

UCLA

UCLA Previously Published Works

Title

EPCAM mutation update: Variants associated with congenital tufting enteropathy and Lynch syndrome.

Permalink

<https://escholarship.org/uc/item/3hx9c0v4>

Journal

Human mutation, 40(2)

ISSN

1059-7794

Authors

Pathak, Sagar J
Mueller, James L
Okamoto, Kevin
et al.

Publication Date

2019-02-01

DOI

10.1002/humu.23688

Peer reviewed

EPCAM mutation update: Variants associated with congenital tufting enteropathy and Lynch syndrome

Sagar J. Pathak^{1,2} | James L. Mueller¹ | Kevin Okamoto¹ | Barun Das¹ |
 Jozef Hertecant³ | Lynn Greenhalgh⁴ | Trevor Cole⁵ | Vered Pinsk⁶ |
 Baruch Yerushalmi⁶ | Odul E. Gurkan⁷ | Michael Yourshaw⁸ | Erick Hernandez⁹ |
 Sandy Oesterreicher¹⁰ | Sandhia Naik¹¹ | Ian R. Sanderson¹¹ | Irene Axelsson¹² |
 Daniel Agardh¹³ | C. Richard Boland¹⁴ | Martin G. Martin¹⁵ |
 Christopher D. Putnam^{14,16*} | Mamata Sivagnanam^{1,2*} 

¹Department of Pediatrics, University of California, San Diego, La Jolla, California²Rady Children's Hospital, San Diego, California³Genetics/Metabolics Service, Tawam Hospital, Al Ain, United Arab Emirates⁴Liverpool Women's NHS Foundation Trust, Liverpool, UK⁵West Midlands Regional Genetics Service and Birmingham Health Partners, Birmingham Women's Hospital, Birmingham, UK⁶Division of Pediatrics, Pediatric Gastroenterology Unit, Soroka University Medical Center and Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel⁷Department of Pediatrics, Gazi University School of Medicine, Ankara, Turkey⁸Department of Human Genetics, University of California, Los Angeles, Los Angeles, California⁹Pediatric Gastroenterology, Miami Children's Health System, Miami, Florida¹⁰Rocky Mountain Health Pediatric Gastroenterology, Lone Tree, Colorado¹¹Paediatric Gastroenterology, Barts and the London School of Medicine, London, UK¹²Department of Pediatrics, Skane University Hospital, Malmo, Sweden¹³Department of Clinical Sciences, Lund University, Skane University Hospital, Malmo, Sweden¹⁴Department of Medicine, University of California, San Diego, La Jolla, California¹⁵Department of Pediatrics, University of California, Los Angeles, Los Angeles, California¹⁶San Diego Branch, Ludwig Institute for Cancer Research, La Jolla, California

Correspondence

Mamata Sivagnanam, Division of Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093.

Email: mengineer@ucsd.edu

Funding information

NIH, Grant/Award Numbers: DK083762, R01 DK107764, T32 DK007202; NIH GM, Grant/Award Number: R01 GM26017; NIH CA, Grant/Award Number: R01 CA72851; California Institute for Regenerative Medicine, Grant/Award Number: RT2-01985; Ludwig Institute for Cancer Research

*These authors contributed equally to this work.

Communicated by Rolf H. Sijmons

Abstract

The epithelial cell adhesion molecule gene (*EPCAM*, previously known as *TACSTD1* or *TROP1*) encodes a membrane-bound protein that is localized to the basolateral membrane of epithelial cells and is overexpressed in some tumors. Biallelic mutations in *EPCAM* cause congenital tufting enteropathy (CTE), which is a rare chronic diarrheal disorder presenting in infancy. Monoallelic deletions of the 3' end of *EPCAM* that silence the downstream gene, *MSH2*, cause a form of Lynch syndrome, which is a cancer predisposition syndrome associated with loss of DNA mismatch repair. Here, we report 13 novel *EPCAM* mutations from 17 CTE patients from two separate centers, review *EPCAM* mutations associated with CTE and Lynch syndrome, and structurally model pathogenic missense mutations. Statistical analyses indicate that the c.499dupC (previously reported as c.498insC) frameshift mutation was associated with more severe treatment regimens and greater mortality in CTE, whereas the c.556-14A>G and c.491+1G>A splice site mutations were not correlated with treatments or outcomes significantly different than random

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2018 The Authors. *Human Mutation* published by Wiley Periodicals, Inc.

simulation. These findings suggest that genotype–phenotype correlations may be useful in contributing to management decisions of CTE patients. Depending on the type and nature of *EPCAM* mutation, one of two unrelated diseases may occur, CTE or Lynch syndrome.

KEYWORDS

congenital tufting enteropathy, *EPCAM*, genotype-phenotype correlation, in silico simulation, Lynch syndrome, protein modeling

1 | BACKGROUND

Epithelial cellular adhesion molecule (*EPCAM*, formerly known as *TACSTD1* or *TROP1*; MIM# 185535) encodes the EpCAM protein, which is normally expressed on the basolateral membrane of cells in epithelial tissues and on plasma cells and was originally identified as a tumor antigen (Bergsagel, Victor-Kobrin, Timblin, Trepel, & Kuehl, 1992; Herlyn, Steplewski, Herlyn, & Koprowski, 1979; Litvinov, Bakker, Gourevitch, Velders, & Warnaar, 1994; Litvinov, Velders, Bakker, Fleuren, & Warnaar, 1994). Mutations in *EPCAM* have been linked to two distinct diseases, congenital tufting enteropathy (CTE) and Lynch syndrome (previously called hereditary nonpolyposis colorectal cancer or HNPCC).

1.1 | *EPCAM* in CTE

Biallelic inactivation of *EPCAM* (Sivagnanam et al., 2008) causes CTE (MIM# 613217), which is a rare autosomal recessive diarrheal disorder (Reifen, Cutz, Griffiths, Ngan, & Sherman, 1994) with an estimated incidence of one in 50,000–100,000 births based on a study in Western Europe (Goulet, Salomon, Ruemmele, de Serres, & Brousse, 2007). In addition to being caused by *EPCAM* mutations (Sivagnanam et al., 2008), mutations in *SPINT2* (MIM# 605124) have been implicated in a syndromic form of the disease (Salomon et al., 2014), which may cause an indirect loss of EpCAM protein due to proteolysis by activation of matriptase (Wu, Feng, Lu, Morimura, & Udey, 2017). CTE presents within the first months of life with severe chronic watery diarrhea and growth restriction. CTE patients require rapid diagnosis and often immediate total parenteral nutrition (TPN) therapy to sustain adequate caloric and fluid intake, and to prevent mortality by effectively bypassing the gut (Beck, Kang, & Suh, 2001; Goulet et al., 2007). This disease most often leads to intestinal failure and lack of enteral autonomy, although favorable outcomes have been reported in some patients (Lemale et al., 2011). The current diagnostic criteria include total or partial villus atrophy and crypt hyperplasia typically without evidence of inflammation, in addition to characteristic focal epithelial tufts composed of enterocytes with plasma membrane rounding found in the duodenum and jejunum (Sherman, Mitchell, & Cutz, 2004). The focal epithelial tufts distinguish CTE from the other two congenital enteropathies directly affecting enterocytes: (1) microvillous inclusion disease (MIM# 251850), which is caused by homozygous or compound heterozygous mutations in *MYO5B* and is characterized by Periodic Acid-Schiff positive inclusion bodies in the context of microvillous atrophy (Cutz et al., 1989), and (2) tricho-hepato-enteric syndrome (MIM#

222470 and MIM# 614602), which is caused by homozygous or heterozygous mutations in *TTC37* or *SKIV2L* and is characterized by facial dysmorphism, hair abnormalities, intrauterine growth restriction, and immune deficiency (Verloes et al., 1997).

1.2 | *EPCAM* in Lynch syndrome

In contrast to the biallelic inactivation of *EPCAM* in CTE, monoallelic deletions of the last exons of *EPCAM* give rise Lynch syndrome in 1–3% of affected families (Kovacs, Papp, Szentirmay, Otto, & Olah, 2009; Ligtenberg et al., 2009). Lynch syndrome (MIM# 120435) is an autosomal dominant disease that predisposes patients to early onset colorectal and endometrial cancers (Lynch et al., 2009). Lynch syndrome is caused by inheritance of one defective allele in a DNA mismatch repair (MMR) gene, which is a pathway that primarily acts to repair errors made during replication (Kolodner, 2016). Inheritance of two defective alleles of a MMR gene, however, does not cause Lynch syndrome but rather causes constitutional MMR deficiency (CMMR-D also called biallelic MMR deficiency and MMR cancer syndrome; MIM# 276300), which is associated with blood-borne cancers, central nervous system cancers, and colorectal tumors usually occurring in the first two decades of life (Shlien et al., 2015). Remarkably, Lynch syndrome is caused by defects in only a few of the genes involved in MMR, predominantly *MSH2* and *MLH1*, although carriers of *PMS2* and *MSH6* mutations also have Lynch syndrome though with a milder phenotype than carriers of *MSH2* or *MLH1* defects (Lynch et al., 2009).

Cancer predisposition in Lynch syndrome patients is due to a secondary somatic inactivation of the functional MMR allele (Hemminki et al., 1994). These tumors have strong mutator phenotype and are clinical recognized by the instability of the number of copies of microsatellite repeats (a microsatellite instability high or MSI-H phenotype). Lynch syndrome was traditionally suspected in patients presenting with colon cancer before age 50, or a family history colon or uterine cancer at a young age. The Amsterdam I and II criteria, and more recently, the Bethesda and the revised Bethesda guidelines, propose to identify individuals who were likely to be mutation carriers for Lynch syndrome (Giardiello et al., 2014). These guidelines identify individuals with colorectal cancer and endometrial cancer who should undergo tumor testing for a MSI-H phenotype and immunohistochemical detection of MMR proteins, as well as testing for mutations in MMR genes (Wimmer & Etzler, 2008; Zeinalian, Hashemzadeh-Chaleshtori, Salehi, & Emami, 2018). Once Lynch syndrome is identified, each case belongs to a family that requires genetic counseling, DNA testing for the detected mutation, and screening for colorectal

cancer. Colonoscopy is the mainstay of colorectal cancer surveillance, particularly given that 70–80% of tumors are found proximal to the splenic flexure. The average age of colorectal cancer due to Lynch syndrome is 45 years, thus screening colonoscopy for Lynch syndrome patients is recommended at age 25 and repeated every 1 to 2 years. If patients are found to have colorectal cancer, subtotal colectomy may be necessary, given the marked frequency of synchronous and metachronous colorectal cancers (Lynch et al., 2009).

EPCAM-associated Lynch syndrome is not due to loss of *EPCAM* per se, but rather is due to monoallelic inactivation of the downstream *MSH2* gene (MIM# 609309) by methylation in the CpG island that harbors the promoter and start site of *MSH2* (Fishel et al., 1993). A causative role for silencing *MSH2* in Lynch syndrome in patients with germline 3' deletions of *EPCAM* has been demonstrated by the fact that second hits in tumors arise from somatic variants that affect the unsilenced copy of *MSH2* and are associated with the loss of the *MSH2* protein (Eguchi et al., 2016; Huth et al., 2012; Kovacs et al., 2009; Ligtenberg et al., 2009; Spaepen et al., 2013). Moreover, the presence of multiple Alu sequences from the 3' end of *EPCAM* to the middle of *MSH2* can lead to internal recombination events and deletions of portions of *MSH2* (Hitchins & Burn, 2011; Perez-Cabornero et al., 2011; van der Klift et al., 2005). Losses of only the polyadenylation signal of *EPCAM*, but not the promoter or start site of *MSH2*, leads to a Lynch syndrome-*MSH2* phenotype in those tissues that express *EPCAM*, and may present clinically as a colon cancer-only phenotype (Lynch et al., 2011).

In summary, depending on the type of mutations of *EPCAM*, biallelic or monoallelic, two very distinct clinical phenotypes may occur.

2 | BIOLOGICAL ROLE OF *EPCAM*

In normal tissues, human *EpCAM* is expressed on the surface of simple, pseudo-stratified, and transitional epithelia in various tissues of the gastrointestinal tract, reproductive system, and respiratory tract. *EpCAM* has been implicated in a variety of important cellular functions including cell proliferation, migration, adhesion, differentiation, and cell signaling (Balzar et al., 1999; Litvinov et al., 1997; Trzpis, McLaughlin, de Leij, & Harmsen, 2007; Tsaktanis et al., 2015; Winter et al., 2003). *EpCAM* was first recognized as an antigen overexpressed on human carcinoma cells, including tumors of the gastrointestinal system, breast, thyroid, and kidney (Balzar, Winter, De Boer, & Litvinov, 1999). Decreased and altered *EpCAM* expression has been described in ulcerative colitis patients where altered intestinal barrier function is a major characteristic (Furth et al., 2006). *EpCAM* expression is tightly regulated during embryogenesis, and spatiotemporal patterning of *EpCAM* has recently been shown to be important for embryonic endodermal and mesodermal differentiation (Sarrach et al., 2018).

EpCAM has long been described as a cell adhesion molecule that mediates cell–cell interactions through homooligomerization (Balzar et al., 1999; Litvinov et al., 1997). In contrast, recent experiments have not been able to demonstrate *EpCAM* protein oligomerization in vitro and have not observed a role for *EpCAM* in the adhesion of carcinoma cell lines (Gaber et al., 2018; Tsaktanis et al., 2015). Additionally,

EpCAM expression weakens cadherin-mediated cell adhesion and promotes cell migration and cell invasion in a breast cancer cell line (Osta et al., 2004; Winter et al., 2003). These more recent observations have led to the hypothesis that *EpCAM* requires a ligand for oligomerization or that *EpCAM* is not a cell adhesion molecule. Intriguingly, *EpCAM* lacks structural similarity with other known cell adhesion molecules, and the closely related homolog encoded by *TACSTD2/TROP2* (MIM# 137290) does not mediate cell–cell interactions (Fornaro et al., 1995; Gaber et al., 2018). Unlike *EPCAM*, homozygous and compound heterozygous mutations in the *TACSTD2* gene do not have an intestinal phenotype, but rather give rise to gelatinous drop-like corneal dystrophy (MIM# 204870; Ren et al., 2002; Tsujikawa et al., 1999).

In spite of the controversy regarding specific role of *EpCAM* in mediating cell–cell by homooligomerization, a physical interaction between *EpCAM* and Claudin-7, which is a component of tight junctions, has been demonstrated (Ladwein et al., 2005; Wu et al., 2017), and *EpCAM* contributes to formation of the intestinal barrier by recruiting claudins to cell–cell junctions (Lei et al., 2012). In normal intestines, Claudin-7 forms a stable protein complex with *EpCAM*, whereas CTE-associated *EPCAM* mutations lead to instability of the *EpCAM*–Claudin-7 interaction, decreasing its presence on the plasma membrane (Mueller, McGeough, Pena, & Sivagnanam, 2014). In addition to the critical role *EpCAM* plays in maintenance of the intestinal barrier, *EpCAM* has potent effects on cell proliferation and differentiation (Tsaktanis et al., 2015). *EpCAM* appears to play a role in morphogenesis and development of some tissues. In renal tissue, *EpCAM* is upregulated postischemia in murine models and thus may have a role in regeneration (Trzpis et al., 2007; Trzpis et al., 2008). In the pancreas, *EpCAM* is involved in islet development (Cirulli, Ricordi, & Hayek, 1995; Vercollone, Balzar, Litvinov, Yang, & Cirulli, 2015); however, its function during gut formation has not yet been elucidated. Despite the ubiquity of *EpCAM* in various tissues, the primary symptoms of CTE patients are due to intestinal defects and other systemic defects are not typically reported. *EpCAM*'s biologic role is multifaceted and fully understanding its part in health and disease remains a topic of ongoing investigation.

3 | VARIANTS OF *EPCAM* IN HUMAN DISEASE

The past decade has seen advances in our understanding of the genetics of *EPCAM*. CTE-causing biallelic *EPCAM* mutations are predicted to disrupt the expression and/or stability of the *EpCAM* protein, whereas Lynch syndrome is caused by monoallelic mutations affecting downstream *MSH2*. We present here novel CTE-causing variants of *EPCAM* and review all known variants in relation to these disease processes.

3.1 | Novel variants of *EPCAM* in CTE

To identify additional CTE-causing *EPCAM* variants, 17 CTE patients were recruited from around the world to the University of California, San Diego (UCSD) and the University of California, Los Angeles (UCLA). Informed consent of the subjects was obtained according to the Institutional Review Board guidelines at UCSD and UCLA. Phenotypic data

TABLE 1 Novel variants of *EPCAM* in CTE patients

Patient**	Coding DNA	Protein	Type of mutation	Zygosity	Ethnicity	Gender	TPN requirement	Transplant
1	c.113G>A; c.48_68del21	p.(C38Y); p.(A18_Q24del)	Missense; in-frame deletion	Compound heterozygous	Caucasian	M	Full TPN	No
2	c.1A>C; c.38_62dup25	p.(M1L); p.(A22Cfs*17)	Missense; frameshift	Compound heterozygous	Palestinian	F	Off TPN due to limited access	No
3	c.267G>C; c.*118T>C	p.(Q89H); unknown consequence	Missense; unknown	Compound heterozygous	English	M	Full TPN	No
4	c.307G>A; c.492-5T>C	p.(G103R); splicing defect	Missense; splicing defect	Compound heterozygous	English-Italian	M	Partial TPN	No
5.1	c.491+1G>A; c.556-14A>G	p.(W143_T164del) [del exon 4]; p.(Y186_D219del*)	Splicing defect; splicing defect	Compound heterozygous	Bangladeshi	M	Full TPN	No
5.2	c.491+1G>A; c.556-14A>G	p.(W143_T164del) [del exon 4]; p.(Y186_D219del*)	Splicing defect; splicing defect	Compound heterozygous	Bangladeshi	M	Full TPN	No
6	c.492-1G>A; c.491+1G>A	p.(W143_T164del) [del exon 4]; p.(W143_T164del) [del exon 4]	Splicing defect; splicing defect	Compound heterozygous	Hispanic	M	Full TPN	No
7	c.509_511del TCA	p.(I170del)	In-frame deletion	Homozygous	Pakistani	Unknown	Unknown	Unknown
8	c.555+1G>C	p.(A65Mfs*23) [del exon 5]	Splicing defect	Homozygous	Turkish	F	Weaned off	No
9.1	c.579delT	p.(I193Mfs*17)	Frameshift	Homozygous	Bedouin	M	Full TPN	No
9.2	c.579delT	p.(I193Mfs*17)	Frameshift	Homozygous	Bedouin	F	Full TPN	No
10	c.589C>T	p.(Q197*)	Truncation	Homozygous	Iraqi	M	Full TPN	No
11.1	c.540delT; c.491+1G>A	p.(F180Lfs*30); p.(W143_T164del) [del exon 4]	Frameshift; splicing defect	Compound heterozygous	Unknown	M	Partial TPN	No
11.2	c.540delT; c.491+1G>A	p.(F180Lfs*30); p.(W143_T164del) [del exon 4]	Frameshift; Splicing defect	Compound heterozygous	Unknown	M	Partial TPN	No
12	c.757G>A	p.(D253N)	Missense	Homozygous	Turkish	M	Full TPN	No
13	g.29411_34526 del5116	p.(Q24_Y142del)	Deletion	Homozygous	Hispanic	F	Partial TPN	No
14	c.227C>G	p.(S76*)	Truncation	Homozygous	Israeli	F	Full TPN	No

**Siblings are annotated with decimals, e.g. patients 5.1 and 5.2.

were ascertained. Exons of the *EPCAM* gene were amplified by polymerase chain reaction from patient genomic DNA and subjected to Sanger sequencing as previously described (Sivagnanam et al., 2008; Sivagnanam et al., 2010).

Of the 17 newly identified patients, there were 11 males, five females, and the gender of one patient was not obtained. The disease outcomes included: 10 patients treated with full TPN, four patients treated with partial TPN, and two patients weaned off TPN (one due to limited vascular access); the treatment of one patient was unknown. No patients underwent bowel transplantation or were deceased. Nine patients were compound heterozygotes and eight were homozygotes for mutations in *EPCAM* (Table 1).

A total of 19 distinct *EPCAM* mutations were observed, and 13 of these mutations have not been previously reported (c.1A>C, c.113G>A, c.227C>G, c.267G>C, c.589C>T, c.757G>A, c.540delT, c.579delT, c.38_62dup25, c.48_68del21, c.509_511delTCA,

c.492-5T>C [NC_000002.12:g.47377009T>C], c.*118T>C, and a deletion spanning the end of exon 1 and intron 3–4, NG_012352.2: g.29411_34526del5116, which generates p.Q24_Y142del; note that coding mutations are reported relative to the *EPCAM* reference sequence NM_002354.2). None of the novel *EPCAM* mutations were observed more than once outside of individual families or in homozygous patients. This pattern is consistent with CTE affecting children of families that carry rare recessive *EPCAM* mutations and does not provide evidence for a founder effect in these patient cohorts, unlike that observed in some patients of Kuwaiti and Middle Eastern origin (Salomon et al., 2011).

3.2 | Reported variants in *EPCAM* in CTE

An analysis of the 72 previously reported CTE patients (AlMahamed & Hammo, 2017; Al-Mayouf, Alswaied, Alkuraya, AlMehaidib, & Faqih,

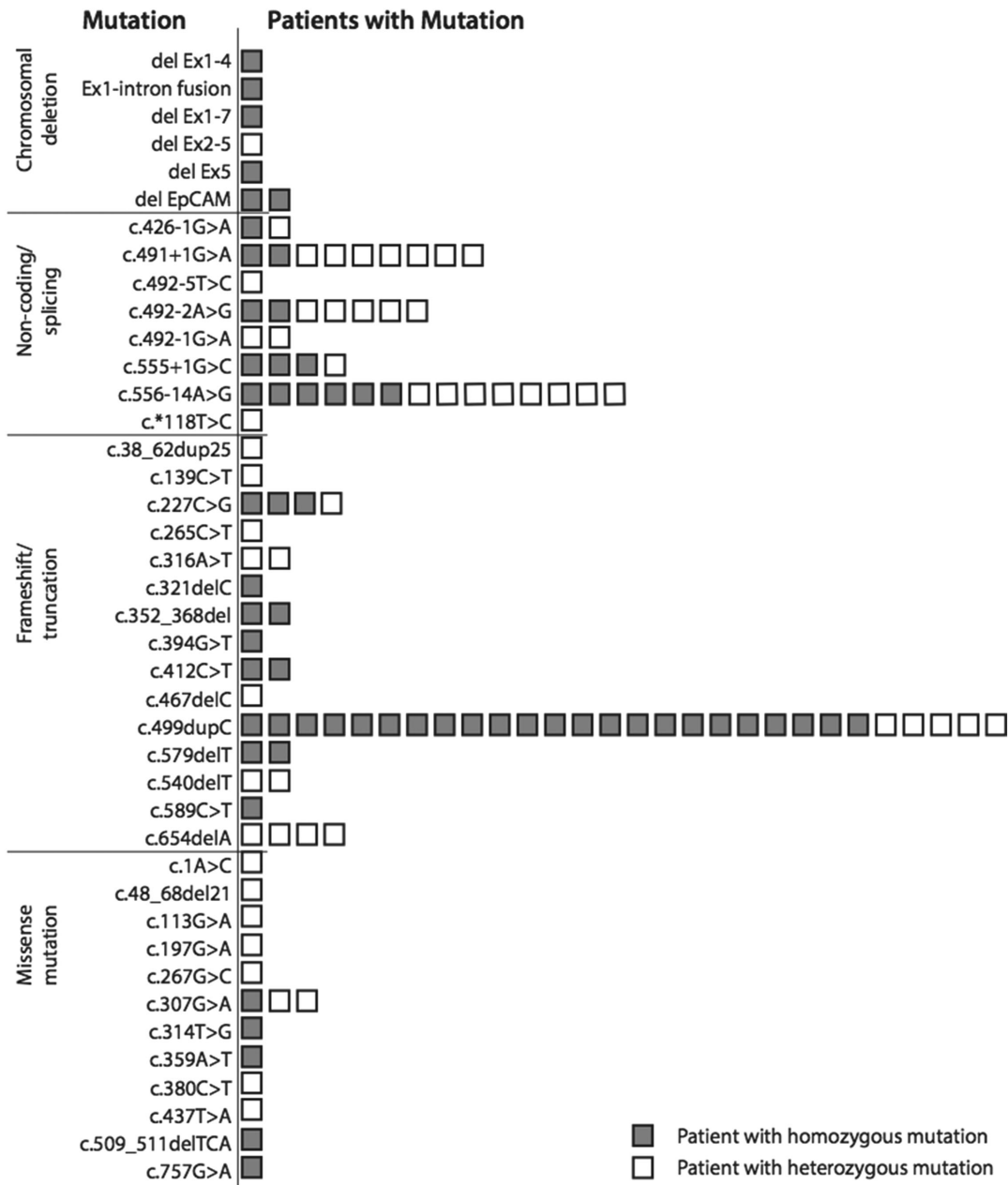


FIGURE 1 Quantitative distribution of *EPCAM* mutations identified in CTE patients. Each patient is represented by a single box. Gray boxes correspond to the presence of the mutation in homozygous patients, whereas white boxes correspond to the presence of a mutation in heterozygous patients

2009; Bodian et al., 2017; d'Apolito et al., 2016; Ko et al., 2010; Pégas et al., 2014; Salomon et al., 2011; Salomon et al., 2014; Schnell et al., 2013; Shakhnovich, Dinwiddie, Hildreth, Attard, & Kingsmore, 2017; Sivagnanam et al., 2008; Sivagnanam et al., 2010; Tang, Huang, Xu, & Huang, 2018; Thoeni et al., 2014) along with the 17 novel patients reported here is consistent with the expected genetics of an autosomal recessive disease (Figure 1; Supporting Information Table 1). Many CTE patients (60 of 90 patients) were homozygous for *EPCAM* mutations, and most of the remainder were compound heterozygotes (24 of 90 patients). The large proportion of homozygote individuals is consistent with the high incidence of CTE patients amongst consanguineous families (21 of the 30 cases where information was available) and/or were consistent with founder effects, as has been previously

reported (Salomon et al., 2011). Additionally, for the 88 patients where gender was known (50 males and 38 females), there was no bias ($P = 0.36$, chi-squared test). The majority of patients are Middle Eastern in descent, though several other ethnicities were represented, including Caucasian and Hispanic.

The 90 reported CTE patients implicate 42 distinct *EPCAM* mutations in causing CTE (Figures 2a and 2b). These mutations include six chromosomal deletions, eight noncoding/splicing mutations, 16 frameshifts/nonsense mutations that would lead to *EpCAM* truncation, and 12 missense mutations or in-frame deletions. Only four mutations have been reported in five or more patients: c.499dupC (previously reported as c.498insC; 28 patients, 23 homozygotes), c.556-14A>G (14 patients, six homozygotes), c.491+1G>A (nine

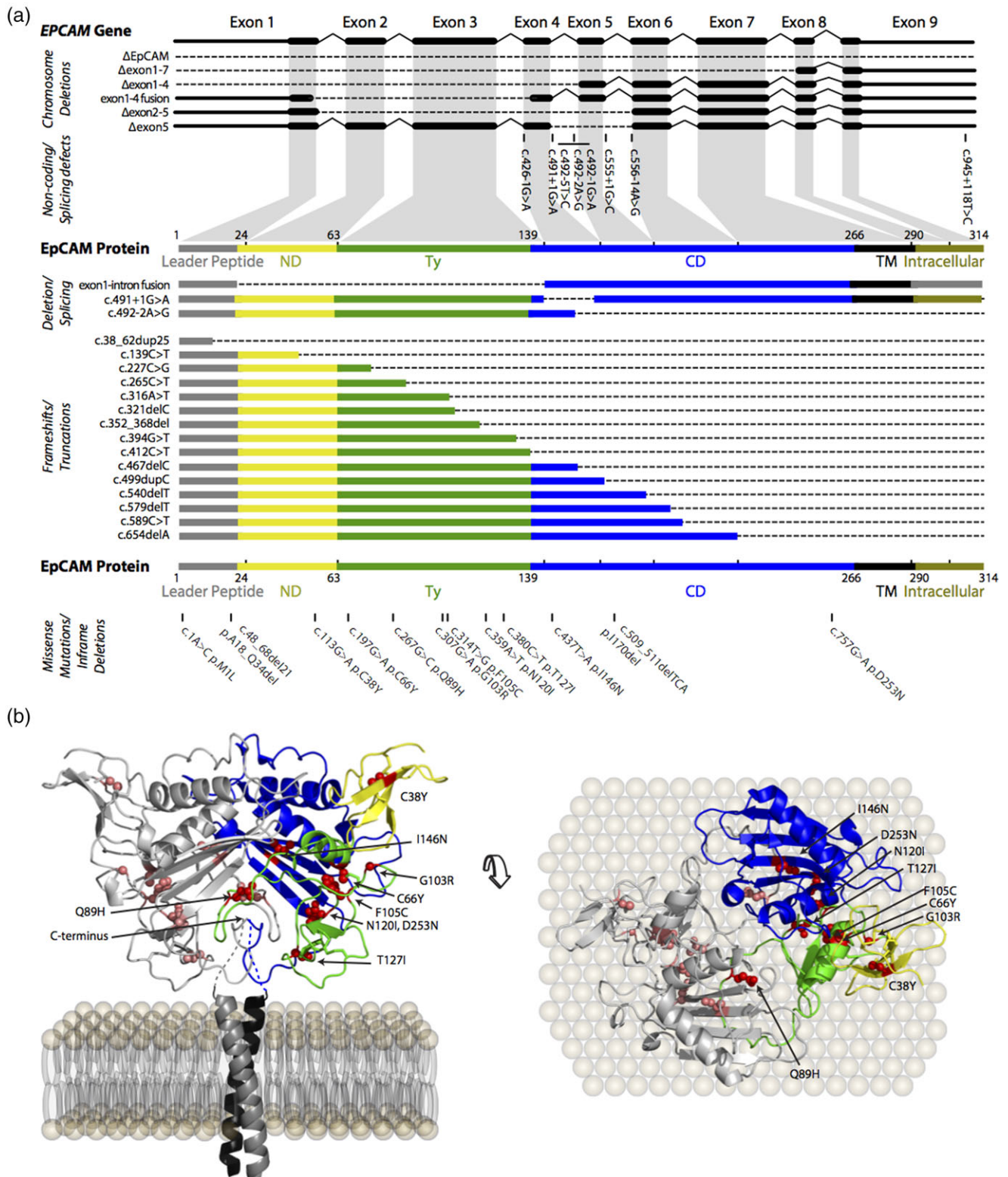


FIGURE 2 Reported mutations in *EPCAM* identified in CTE patients. (a) Reported mutations are shown based on their effect on the gene (top), mRNA (middle), or protein (bottom). (b) The positions of missense mutations are depicted on the EPCAM dimer structure; theoretical membrane and transmembrane helix structures are shown to indicate protein orientation

patients, two homozygotes), and c.492-2A>G (seven patients, two homozygotes).

3.3 | Reported variants of *EPCAM* in Lynch syndrome

The *EPCAM* gene is 17 kb upstream of the *MSH2* gene on the short arm of human chromosome 2. The first evidence that large deletions upstream of the *MSH2* gene could give rise to Lynch syndrome was reported in families from Swiss and American populations (van der Klift et al., 2005). Around the same time, heritable hypermethylation of the *MSH2* promoter, which was the highest in colonic mucosa and colon tumors but lowest in blood leukocytes, was reported (Chan et al., 2006). These two observations turned out to be linked, as monoallelic deletions of the 3' end of the *EPCAM* gene in which the polyadenylation signal is lost give rise to *MSH2* promoter hypermethylation, read-through transcription of the *EPCAM* and *MSH2* genes, and loss of *MSH2* protein expression (Kovacs et al., 2009; Ligtenberg et al., 2009). As these 3' *EPCAM* deletions are germline variants, the promoter methylation is heritable and has been termed an "epimutation," which has also been associated with the *MLH1* promoter in at least one case of Lynch syndrome (Gazzoli, Loda, Garber, Syngal, & Kolodner, 2002). In some of the observed *EPCAM*-*MSH2* fusion transcripts, a cryptic exon in the intergenic region between *EPCAM* and *MSH2* was present in the transcript (Gazzoli et al., 2002; Kovacs et al., 2009; Ligtenberg et al., 2009).

Twenty-five 3' deletions of *EPCAM* characterized at the sequence level have been implicated in causing Lynch syndrome (Dymerska et al., 2017; Eguchi et al., 2016; Guarinos et al., 2010; Huth et al., 2012; Kempers et al., 2011; Kovacs et al., 2009; Kuiper et al., 2011; Ligtenberg et al., 2009; Lynch et al., 2011; Mur et al., 2014; Nagasaka et al., 2010; Niessen et al., 2009; Perez-Cabornero et al., 2011; Rossi et al., 2017; Rumilla et al., 2011; Spaepen et al., 2013; van der Klift et al., 2005). Most of these deletions are mediated by recombination between imperfectly homologous *Alu* repeats (Figure 3a), which are a class of short repeats (~300 bp) present in >10⁶ copies in the human genome (Deininger, 2011). The length of the microhomologies at the breakpoint junction has been emphasized previously (Kuiper et al., 2011; van der Klift et al., 2005); however, even the shorter microhomology junctions are in the context of the larger *Alu* repeat homology (Figure 3a). Consistent with this, *Alu* sequences on the + strand only recombine with other *Alu* sequences on + strand, and the same is true for *Alu* sequences on the - strand.

Twenty-three of the characterized deletions span *EPCAM* exons 8 and 9 (Figure 3b; Supporting Information Table 2), which eliminates the 3' polyadenylation signal and thereby leads to read-through transcription of *MSH2* and hypermethylation of the *MSH2* promoter. The majority of the deletions only span exons 8 and 9 (Figure 3b), which includes the founder mutation in the Dutch population, *EPCAM* c.859-1452_*1999del, (Ligtenberg et al., 2009; Lynch et al., 2011), a potential founder mutation in the Spanish population, *EPCAM* c.858+2568_*4596del (Guarinos et al., 2010; Mur et al., 2014), and a potential founder mutation in the Polish population, *EPCAM* c.858+2478_*4507del (Dymerska et al., 2017). In principal, only deletion of exon 9 is required to give rise to read-through transcription of

MSH2; however, no *Alu* repeats are present within intron 8 (Figure 3b), suggesting that the common codeletion of exons 8 and 9 is influenced not by a requirement for loss of exon 8 but by the genomic context that leads to *Alu*-*Alu* recombination events. In addition to deletions that only affect the 3' end of *EPCAM*, two deletions spanning the 3' end of *EPCAM* and the 5' end of *MSH2* have also been identified (Perez-Cabornero et al., 2011; Sekine et al., 2017).

In addition to the disease-causing variants reported here, some nonpathogenic variants have been reported in the HGMD (<http://www.hgmd.cf.ac.uk/ac/gene.php?gene=EPCAM>) and Leiden Open Variant Database (LOVD) and are summarized in Supp. Table 3.

4 | VARIANT DATABASE

A locus-specific database in the LOVD for *EpCAM* was created in 2009 and is moderated by Joh Paul Plazzer, Johan den Dunnen, and Mamata Sivagnanam (<https://databases.lovd.nl/shared/variants/EPCAM>). All previously published and novel mutations were recorded after performing a literature search for patients with CTE or Lynch syndrome in whom a variant in the *EPCAM* gene was reported. *EPCAM* genotypes were recorded from the original publications when data were available. Adjustment of nomenclature of certain variations were made to ascertain that base pair and protein sequence numbering were based on the *EPCAM* reference sequence NM_002354.2.

5 | *EPCAM* STRUCTURE AND FUNCTION

The nine exons of the *EPCAM* gene encode a 314 amino acid transmembrane glycoprotein (Figure 2a). The extracellular portion of *EpCAM* and the signal peptide is encoded by exons 1 to 6 and is composed of an N-terminal disulfide-stabilized domain, a central thyroglobulin-like domain, and a C-terminal domain (Pavsic, Guncar, Djinnovic-Carugo, & Lenarcic, 2014). The N-terminal disulfide-stabilized domain has previously been described as an "EGF-like" domain, as the domain was predicted to contain beta-strands and disulfides; however, the structure reveals that it has a disulfide bonding pattern more similar to the Cripto/Fri-1/Cryptic domain of mouse Cripto and a fold more similar to WW domains, which was named due to two conserved tryptophan residues (Pavsic et al., 2014). We therefore term this domain the N-terminal domain in this review, consistent with the terminology in the report of the *EpCAM* crystal structure (Pavsic et al., 2014). The thyroglobulin-like domain has also been referred to as a second "EGF-like" domain; however, the similarity with thyroglobulin was recognized some time ago and verified by the *EpCAM* structure (Baeuerle & Gires, 2007; Pavsic et al., 2014). The transmembrane portion is coded by exon 7, and the intracellular domain is encoded by exons 8 and 9 (last 13 amino acids; Balzar et al., 1999).

5.1 | Predicted structural effects of *EPCAM* variants in CTE

In general, the *EPCAM* mutations found in CTE patients appear to be broadly inactivating. The large chromosomal truncations (6 of

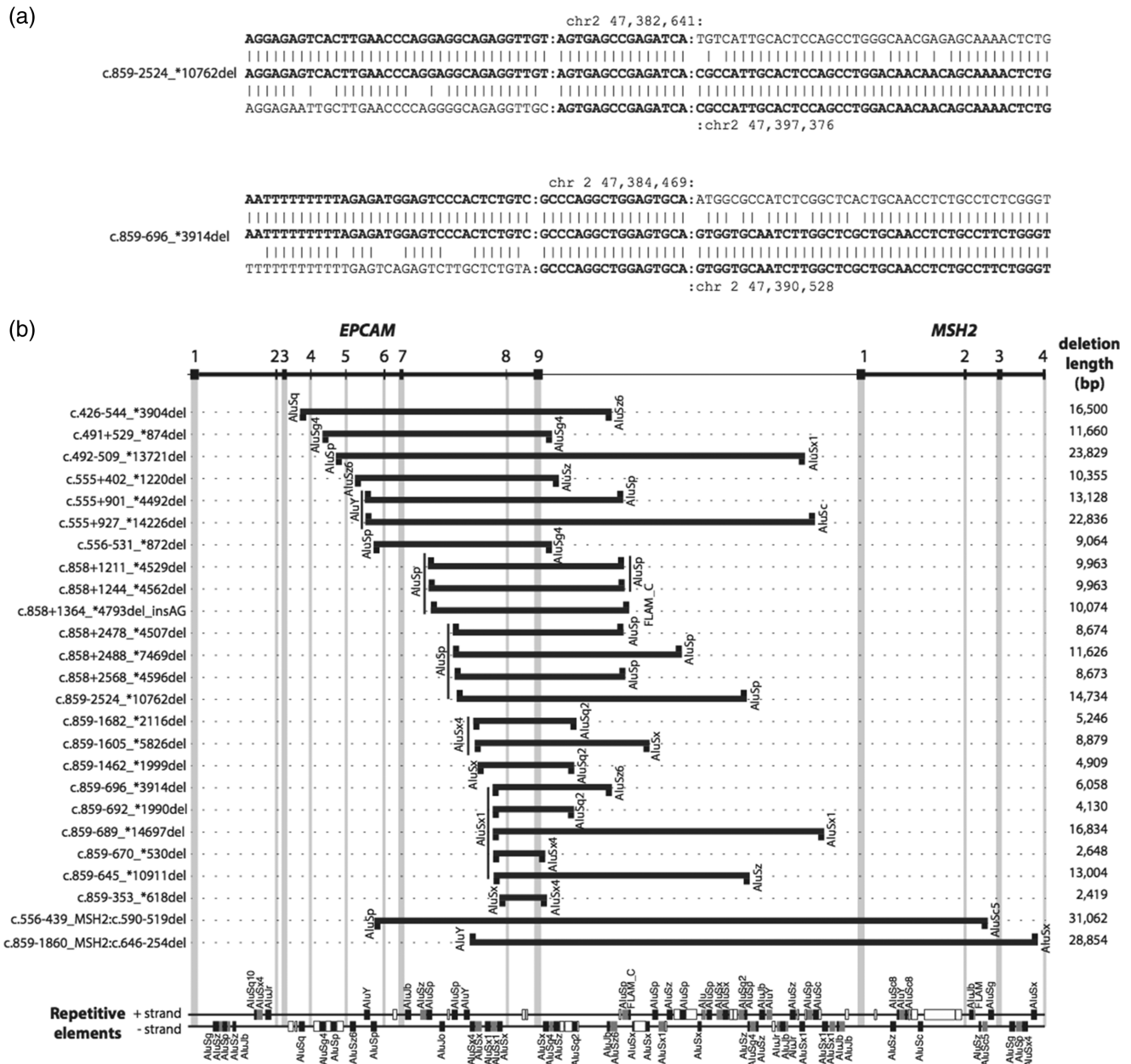


FIGURE 3 3' deletions of *EPCAM* associated with Lynch syndrome. (a) Illustration of the homologies of two *Alu*-*Alu* recombination events causing deletions of *EPCAM* exons 8 and 9. The top and bottom sequences correspond to the two *Alu* sequences involved in the recombination event, and the middle sequence is the novel junction. The sequences between the colons correspond to the identical microhomology at the deletion junction. (b) Diagram of the *EPCAM* deletions. Exons 1-9 of *EPCAM* and the first four exons of *MSH2* are depicted at the top. Deletions are shown in the middle as thick black lines. The *Alu* repeats involved at the homology-mediated rearrangements for each deletion are shown as labeled black boxes; black boxes above the deletion line are *Alu* sequences encoded on the + strand, whereas black boxes below the deletion line are *Alu* sequences encoded on the - strand. Repetitive elements on the + and - strand identified by RepeatMasker version 406 (Smit, 2013-2015) with the Repeat Library Version 20150807 are shown at bottom. *Alu*-related sequences are shown as grey and black boxes; other repeats are shown as white boxes. Note that some deletions have previously been reported under different annotations: c.426-544_*3904del (c.423-545_*3903del), c.555+402_*1220del (AC079775.6:g.72468_82822del10355), c.555+927_*14226del (c.555+894_*14194del), c.858+1211_*4529del (c.85811211_4529del), c.858+1358_*4793del_insAG (c.858+1364_*4793del_insAG), c.858+2478_*4507del (AC079775.6:g.77436_86109del8674), c.859-2524_*10762del (AC079775.6:g.77631_92364del14734), c.859-1605_*5826del (c.859-1605_*5862del), c.859-696_*3914del (AC079775.6:g.79459_85516del6058), c.859-692_*1990del (c.859-672_*2170del), c.859-689_*14697del (AC079775.6:g.79465_96299del16834), and c.859-645_*10911del (AC079775.6:g.79509_92513del13004). In addition, two deletions reported as annotations but not with sequence do not support the reported junction microhomologies: c.492-509_*13721del and c.858-353_*618del

41 mutations) give rise to variants that primarily lack most of the N-terminal extracellular region of EpCAM, whereas the frameshift and nonsense mutations (15 of 41 mutations) truncate the EpCAM prior to the transmembrane helix and C-terminal intracellular region (Pavsic et al., 2014). A small number of mutations were short (<30 bp) duplications or deletions (four of 41) or larger chromosomal deletions (six of 41). Three of these are in-frame deletions: c.48_68del21 (p.A18_Q24del) disrupts the leader peptide targeting EpCAM to the plasma membrane; the c.509_511delTCA deletes I170, which forms part of the hydrophobic core of the C-terminal domain; and the exon 1–4 fusion deletes the N-terminal and thyroglobulin domains. In contrast, the majority of point mutations (31 of 41) were single base insertions, deletions, or changes, and most of these (21 of 31) gave rise to frameshifts, stop codons, or splicing defects. The impact of the observed mutations on the EpCAM protein has not yet been extensively studied, although c.491+1G>A generates mRNA molecules lacking exon 4 and c.492-2A>G causes abnormal in-frame skipping of exon five (Salomon et al., 2011; Sivagnanam et al., 2008).

Because the CTE-associated deletion, frameshift, and splice site mutations in *EPCAM* appear to be loss of function mutations, we aimed to understand how the known amino acid substitution mutations could affect the EpCAM protein structure, as most of these have not been investigated experimentally. The nine *EPCAM* missense mutations encode single amino acid substitutions primarily affected the extracellular thyroglobulin homology domain and internal portions of the extracellular C-terminal domain of EpCAM (Figures 2a and 2b), consistent with the hypothesis that defects in cell–cell interactions are causal in CTE. Several of the amino acid substitutions are poised to disrupt the disulfide bonding in EpCAM extracellular domains (Figure 2b). The c.113G>A (p.C38Y) and c.197G>A (p.C66Y) mutations eliminate cysteine residues involved directly at the disulfide bonds in the N-terminal domain and the thyroglobulin homology domain (Figures 4a and 4b). Additionally, the c.314T>G (p.F105C) mutation provides a competing cysteine that could isomerize the C66–C99 disulfide bond (Figure 4a). The c.307G>A (p.G103R) mutation amino acid substitution is positioned to sterically disrupt the C66–C99 disulfide bond formation (Figure 4a). Other amino acid substitutions are positioned to locally disrupt aspects of the EpCAM structure: the c.437T>A (p.I146N) mutation places a polar residue in a highly hydrophobic environment (Figure 4c), the c.359A>T (p.N120I) and c.757G>A (p.D253N) mutations affect residues that are at one of the two contact regions between the thyroglobulin homology domain and the C-terminal domain (Figure 4d), and the c.380C>T (p.T127I) mutation places a hydrophobic residue in a polar environment and would, at a minimum, disrupt the hydrogen bonding interactions maintained by the amino acid side chains in the wild-type protein (Figure 4e).

5.2 | Predicted structural effects of *EPCAM* variants in Lynch syndrome

Most of the Lynch syndrome associated 3' deletions of *EPCAM* involve deletions of exon 8 and 9. Based on the EpCAM protein structure, these deletions would be predicted to give rise to loss of intracellular portions of EpCAM while leaving the extracellular domains and most of

the transmembrane domain intact (Figure 2a). However, it is unclear if these truncated EpCAM proteins are made, as read-through transcription would also give rise to *EPCAM-MSH2* fusions (often between *EPCAM* exon 7 and *MSH2* exon 2, although an *EPCAM* exon 5-*MSH2* exon 2 fusion was observed with c.555+927_*14226del and c.556-531_*872del) and *EPCAM*-cryptic exon-*MSH2* fusions (Eguchi et al., 2016; Kovacs et al., 2009; Ligtenberg et al., 2009; Perez-Cabornero et al., 2011; Spaepen et al., 2013). Most of the *EPCAM* exon 7-*MSH2* exon 2 fusions and *EPCAM* exon 7-cryptic exon-*MSH2* exon 2 fusions lead to a premature stop in *MSH2* exon 2, likely leading to loss of the mRNA through nonsense-mediated decay (Maquat, 2004). In contrast, some transcripts isolated from the c.859-696_*3914del background involve an alternative splice site donor in *EPCAM* exon 7 and alternative splice site acceptors in *MSH2* exon 2, and at least one transcript leads to an in-frame fusion (Kovacs et al., 2009). In the case of one deletion spanning *EPCAM* exons 6–9 and *MSH2* exons 1–2, a mislocalized cytoplasmic *MSH2* protein was detected by immunohistochemistry (Sekine et al., 2017). Consistent with the possibility that many 3' deletions of *EPCAM* do not produce truncated EpCAM proteins, tumors arising in *EPCAM*-associated Lynch syndrome patients with a secondary loss of the wild-type copy of *EPCAM* do not express EpCAM, as observed by immunohistochemistry (Huth et al., 2012; Kang et al., 2015; Kloor et al., 2011; Musulen et al., 2013). This study, however, does not confirm that loss of only exons 8 and 9 of *EPCAM* is inactivating, as the deletions were not mapped and the multiplex ligation-dependent probe amplification (MLPA) protocol performed only tested exons 3, 8, and 9 of *EPCAM* for copy number changes. In contrast, a number of the larger 3' deletions of *EPCAM* associated with Lynch syndrome (Figure 3b) would give rise to EpCAM truncations that overlap with those observed in CTE and hence can be reliably predicted to be inactivating.

Structural effects of *EPCAM* variants are diverse in nature and lead to the unique phenotypes of CTE versus Lynch syndrome.

6 | ANIMAL MODELS

EpCAM is highly conserved in mammals (Balzar et al., 1999; Bergsagel et al., 1992). The pattern of murine *Epcam* mRNA expression is similar to that of human *EPCAM*, with highest expression in the gut and lower levels in the kidneys, pancreas, mammary glands, thyroid, pituitary, salivary glands, lungs, and genitalia, consistent with its epithelial distribution (Nagao et al., 2009). EpCAM protein has also been identified on the surface of mouse embryonic, neonatal, and adult germ cells, as well as embryonic stem cells (Basak et al., 1998; Gonzalez, Denzel, Mack, Conrad, & Gires, 2009). In vivo roles for EpCAM have been investigated in two zebrafish knockout models, which demonstrated its function in epithelial morphogenesis and skin development (Slanchev et al., 2009; Villablanca et al., 2006). *Epcam* knockout was initially reported to cause embryonic lethality in mice (Nagao et al., 2009); however, three groups have reported viable *Epcam* knockouts (Gaiser et al. 2012; Guerra et al., 2012; Lei et al., 2012). Constitutive deletion of *Epcam* exons 2 and 3 caused intestinal defects and an inability of the mutant mice to gain weight and was associated with a failure to form functional tight

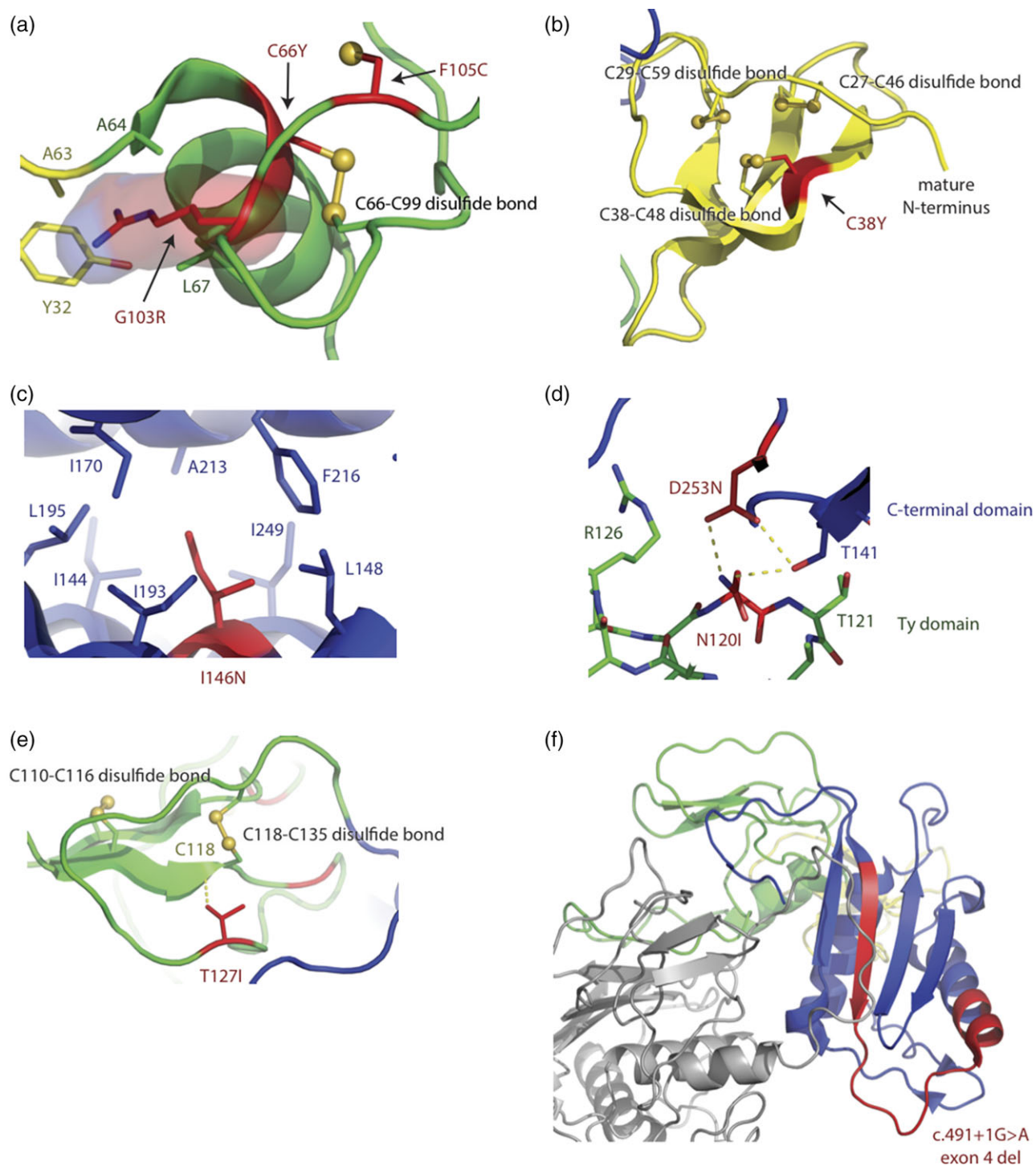


FIGURE 4 Detailed views of the positions of reported missense mutations in *EPCAM* and c.491+1G>A. (a) The F105C and G103R amino acid substitutions are positioned to affect the C66–C99 disulfide bond through disulfide bond isomerization (F105C) or through steric interactions (G103R) as shown by the molecular surface of the arginine mutant that clashes with Y32. (b) The C38Y amino acid substitution disrupts one of the disulfides in the N-terminal domain. (c) The I146N amino acid substitution places a polar residue within the hydrophobic core of the C-terminal domain. (d) The N120I and D253N substitutions disrupt the side chain–main chain hydrogen bonds between the thyroglobulin-like domain and the C-terminal domain. (e) The T127I amino acid substitution would disrupt side chain–main chain interactions adjacent to the C118–C135 sulfide bond in the thyroglobulin-like domain. (f) The c.491+1G>A splice site mutation that causes exon 4 skipping would delete a core region (red cartoon) of the C-terminal domain

junctions and defects in the recruitment of claudins in line with EpCAM's role of recruiting claudins to tight junctions in cellular models (Lei et al., 2012; Wu, Mannan, Lu, & Udey, 2013). Another viable knockout constructed by gene trapping revealed severe hemorrhagic enteropathy (Guerra et al., 2012). EpCAM mice models with neonatal intestinal abnormalities have suggested EpCAM affects cell-cell junctions through expression and localization of E-cadherin and beta-catenin, which are essential components of adherens junctions (Guerra et al., 2012), and decreased expression of tight junction proteins, increased permeability, and decreased ion transport in the intestines (Kozan et al., 2015). Additionally, a conditional knockout of *Epcam* in Langerhans cells demonstrated a role for EpCAM in promoting the motility and migration of these cells from the skin to the lymph nodes after activation (Gaiser et al., 2012). Constitutive and inducible CTE-associated murine models express mutant EpCAM, lacking exon 4, rather than a gene knockout, and mimic CTE with growth retardation, and intestinal epithelial tufts. These models show enhanced intestinal permeability and migration as well as decreased ion transport and expression of tight junctional proteins (Kozan et al., 2015; Mueller et al., 2014).

7 | GENOTYPE-PHENOTYPE CORRELATION

The dramatically different phenotypes of CTE and Lynch syndrome provide a clear genotype-phenotype correlation: monoallelic 3' deletion of *EPCAM* leading to loss of *MSH2* expression does not cause CTE, but does give rise to Lynch syndrome, whereas biallelic inactivation of *EPCAM* gives rise to CTE. As CTE is caused directly by loss-of-function *EPCAM* mutations, a large number of inactivating mutations can cause CTE without affecting expression of *MSH2* and hence do not cause Lynch syndrome or CMMR-D. In principle, genotypes can be envisioned that combine both CTE and MMR defects. A combination of Lynch syndrome and CTE could theoretically arise in a heterozygote possessing a 3' deletion of *EPCAM* that inactivates *MSH2* and another *EPCAM* inactivating mutation that does not inactivate *MSH2*. A combination of CMMR-D and CTE could arise by germline biallelic inactivation of *EPCAM* via 3' deletions that also lead to loss of *MSH2* expression. To date, no patient has been described with both CTE and Lynch syndrome or CMMR-D, which may be due in part to the fact that Lynch syndrome is observed at an age past the typical lifespan of CTE patients. However, a 9-year-old patient with a heterozygous deletion of *EPCAM* and a heterozygous *MSH2* missense mutation was identified as having atypical CMMR-D in which *MSH2* was absent in an EpCAM expression-dependent manner (Li-Chang et al., 2013). This patient was not reported to have CTE, likely due to EpCAM expression from the normal allele.

7.1 | Genotype-phenotype correlations in CTE

A wide range of alimentary needs were noted among the 68 reported CTE patients for whom nutritional data were available (Figure 5): four patients were weaned off TPN entirely; 16 patients had partial TPN ranging from 3 out of 7 (3/7) days to 6/7 days; and 48 patients were

on full TPN (7/7 days). About one-third of patients showed significant morbidity and mortality as 13 patients underwent transplant and 12 patients were deceased. Of the transplanted patients, four patients were deceased. We hypothesized these differential clinical outcomes, nutritional data/TPN status, need for bowel transplantation, and mortality, could be used as surrogate markers for disease severity. To examine these data for evidence of a genotype-phenotype correlation, patients were divided into genotypic groups in which *EPCAM* alleles were mapped to categories (frameshift mutation, nonsense mutation, missense mutation, and splicing defect). The nutritional requirements, need for bowel transplantation, and mortality for each genotypic group were then compared to the rest of the patients using the maximum likelihood G-test (Supporting Information Table 3), which is more appropriate than the chi-squared test for small sample sizes (Sokal, 1994). The only genotypic group with significant differences after multiple-test correction with a false discovery rate of 0.05 was the frameshift/frameshift genotypic group; these patients were more likely to require full TPN (16 of 16 patients; uncorrected *P*-value 0.000347; G-test) and be deceased (eight of 16 patients; uncorrected *P*-value 0.000473; G-test). Thus, these patients are associated with worse clinical outcomes and potentially more severe disease.

The small number of characterized patients and the presence of relatively large numbers of compound heterozygotes complicate the use of statistical tests like the G-test and chi-squared test for correlating individual mutations with disease severity. We therefore performed random (Monte Carlo) permutation tests *in silico* to look for individual classes of mutations (frameshift, nonsense, missense, and splice defect) in genotypes associated with the surrogate markers of disease severity. These analyses assume that mutations that would cause more severe disease in homozygote patients will tend to be somewhat more deleterious in compound heterozygotes and that mutations causing less severe disease will also influence severity in compound heterozygotes. In each trial of this simulation, the 68 nutritional (weaned, partial TPN, or full TPN), transplant (yes or no), and mortality (yes or no) outcomes were randomly distributed among the 68 patient genotypes. In other words, the clinical outcomes were randomly permuted in each round; permutation tests are useful nonparametric methods of hypothesis testing when null hypotheses are difficult or impossible to obtain (Fisher, 1935; Pitman, 1937; Pitman, 1938). These random trials were run 10 million times. For each nutritional, transplant, and mortality outcome, a distribution of the number of times the mutation of interest was associated with the outcome was generated. These distributions are null distributions that assumed the clinical outcomes were independent of the mutations in the genotypes. These null distributions were then compared to the number of associations of clinical outcomes to mutations in the real data. *P*-values were calculated by summing the probabilities of all events in the random distribution that were equal to or more extreme than the number of associations in the real data. This simulation strategy accounted for the relative frequencies of the mutations as well as the nonrandom cooccurrence of mutations in the observed patient genotypes. Consistent with the statistical test results for the frameshift/frameshift genotypes, frameshift mutations were more frequently present in genotypes of patients requiring full TPN than predicted by chance (*P* = 0.003; Figure 6a) and were

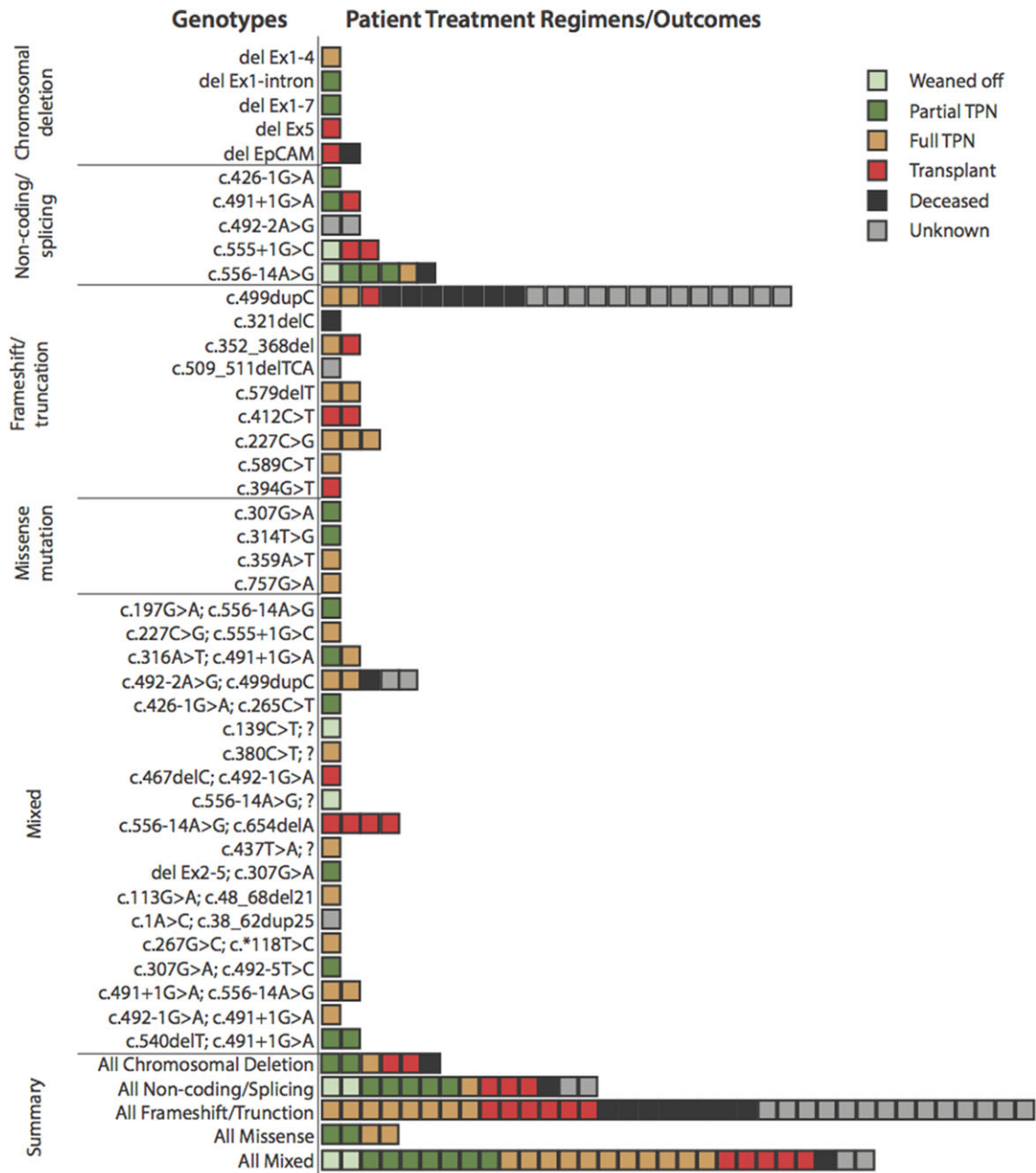


FIGURE 5 Quantitative distribution genotypes in the current literature. Colors denote current treatment status or clinical outcome of patients with the indicated genotype; each patient is represented by a single box. Transplanted patients who were deceased are reported as transplanted

less frequent in genotypes of patients with partial TPN ($P = 0.02$; Figure 6a). Frameshift mutations were also more commonly present in patients who were deceased ($P = 0.002$; Figure 6a). In contrast, splice site mutations were less commonly present in patients requiring full TPN than predicted by chance ($P = 0.007$): more commonly present in patients with partial TPN ($P = 0.03$; Figure 6a). Neither nonsense mutations nor missense mutations had a statistically significant association with any clinical outcome (Supporting Information Figure 1).

We then used this Monte Carlo simulation strategy to analyze individual mutations. The c.499dupC mutation is present in 13 of the 69 patients for which severity information was available. All 13 patients with c.499dupC were treated with full TPN, which is significantly

more TPN treatment relative to the random simulation ($P = 0.006$) and is significantly less partial TPN relative to the random simulation ($P = 0.019$; Figure 6b). Additionally, the c.499dupC mutation has a significant increase in association with mortality relative to the random simulation (eight of the 13 patients were deceased; $P = 0.0006$). No significant differences were observed for weaning from full TPN ($P = 0.41$) or bowel transplantation ($P = 0.13$) relative to the simulation. Taken together, these data suggest that the c.499dupC is correlated with more aggressive treatment and poorer outcomes, consistent with more severe disease. The c.499dupC frameshift mutation is responsible for the effect of the frameshift class of mutations in the simulation (Figure 6a) as well as the correlation of frameshift/frameshift genotype with more severe clinical outcomes (10 of 16 frameshift/frameshift

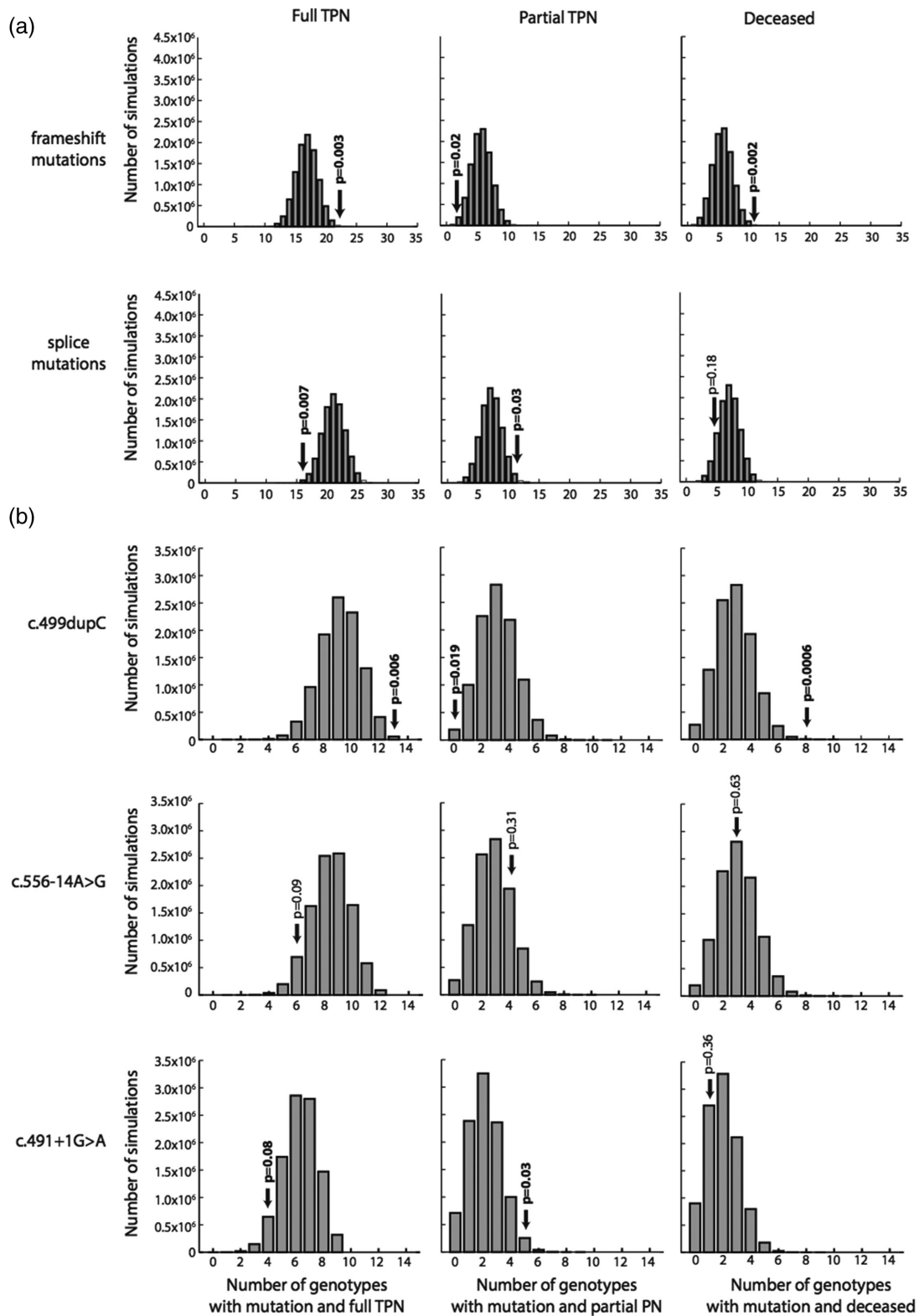


FIGURE 6 Computer simulation of patient treatment/outcomes suggests a correlation between c.499dupC and more severe disease. (a) The number of times a category of mutation was associated with a particular treatment or outcome (vertical arrow) was compared to the expected distribution based on random simulations (grey bars). Reported *P*-values are derived from the expected distributions. Statistically significant *P*-values are in bold. (b) Analysis of individual mutations using random simulation as in panel A. Of the three mutations that could be analyzed, only the c.499dupC mutation had an observed count that was statistically different than the random distribution

genotypes are homozygous c.499dupC mutations; Supporting Information Table 3). In contrast, the c.556-14A>G mutation (present in 14 of the 69 patients) is not correlated with treatments or outcomes that are significantly different than random simulation for full TPN ($P = 0.09$), partial TPN ($P = 0.31$), weaning ($P = 0.14$), transplant ($P = 0.81$), or mortality ($P = 0.63$) (Figure 6b). Similarly, the c.491+1G>A mutation (present in nine of the 69 patients) is also not correlated with significant changes relative to the random simulation for full TPN ($P = 0.08$), partial TPN ($P = 0.03$), weaning ($P = 0.55$), transplant ($P = 0.31$), or mortality ($P = 0.36$) (Figure 6b). Despite the evidence for frameshift mutations in general, and c.499dupC in specific as being associated with more severe clinical outcomes, establishing clear genotype/phenotype correlations in CTE is complicated by several factors. These factors including the rare occurrence of CTE, the fact that the mutations are often present in compound heterozygotes, and the lack of standard quantitative measures for disease severity. Use of clinical treatment, nutritional status, and outcomes to characterize disease is also confounded by the individual circumstances of the patient as well as the resources, experiences, and biases of the treatment centers. The limitations of these analyses argue that the development of a quantitative measure for CTE severity will be necessary to establish genotype-phenotype correlations and inform patient treatment and prognosis.

7.2 | Genotype–phenotype correlations in Lynch syndrome

In *EPCAM*-associated Lynch syndrome, epigenetic silencing of *MSH2* is tissue specific, giving rise to mosaic inactivation of *MSH2*, the high risk of colorectal cancer, and the low risk of endometrial cancer (Dymerska et al., 2017; Kempers et al., 2011; Lynch et al., 2011; Perez-Cabornero et al., 2011). In contrast, deletions that span both *EPCAM* and *MSH2* or span regions close to the *MSH2* promoter give rise to high risk for both colorectal and endometrial cancers like mutations in *MSH2* alone (Kempers et al., 2011).

8 | FUTURE PROSPECTS

EPCAM is involved in different ways in three clinically relevant diseases.

8.1 | Congenital tufting enteropathy

The fact that CTE appears to be a disease caused by relative loss-of-function mutations places constraints on the strategies that can be adopted for long-term therapy that weans patients from TPN. If *EpCAM* functioned to suppress another competing mechanism, then pharmacological intervention of the competing mechanism could be pursued, such as use of matriptase inhibitors in the proposed SPINT2/matriptase mechanism (Wu et al., 2017). Bowel transplant has been used successfully (Paramesh et al., 2003); however, transplants are complicated by the availability of donors, the need to prevent transplant rejection, and the high rate of posttransplantation mortality. Though technology is still evolving, intestinal stem cell transplantation may be the next generation of transplant strategy for

CTE and other monogenic intestinal failure (Hong, Dunn, Stelzner, & Martin, 2017). Mutations affecting *EpCAM* mRNA splicing are relatively common in the population. Some strategies for restoration of aberrant splicing in Duchenne muscular dystrophy and spinal muscular atrophy (Benchouir, Robin, & Goyenvall, 2015; Hua et al., 2010; Hua, Vickers, Okunola, Bennett, & Krainer, 2008; Kole, Krainer, & Altman, 2012; Wu et al., 2017) might be applicable in a subset of CTE patients. A more broadly useful strategy may be to use gene therapy to correct the single nucleotide insertion, deletions, and changes that dominate the CTE patient population. In this regard, the fact that gene therapy may only need to act on cells in the intestinal epithelium may prove to be an advantage; however, it remains an open question as to the level of gene correction that would be required to ameliorate CTE symptoms.

Even in the absence of molecular therapies for CTE, a more pressing concern is the timely diagnosis of infants in order to initiate appropriate treatments as soon as possible. The difficult and lengthy process of histological diagnosis, often requiring repeated endoscopies, is less than ideal. Despite recent immunohistochemical staining advances, including gastric and colonic biopsies when small bowel biopsies are inconclusive, the accuracy and speed remain a limitation of this diagnostic modality (Ranganathan, Schmitt, & Sindhi, 2014; Treetipsatit & Hazard, 2014). Conversely, modern genetic analyses have the promise of accelerating the diagnostic pathway and distinguishing CTE patients from those suffering from other congenital enteropathies (Thiagarajah et al., 2018).

Modern genetic testing methodologies have the advantage of rapidly identifying known mutations and will be particularly useful for the commonly observed mutations. However, the fact that 13 novel mutations were identified in 17 sequenced patients, many of whom were from families without consanguinity, argues that relying solely on previously identified *EPCAM* mutations for diagnosis could lead to many false negatives. Thus, genetic strategies should focus on sequencing of the *EPCAM* gene, potentially using modern short-read DNA sequencing combined with exon-capture technologies, and analysis of exon copy number to identify deletions, potentially using read depth from DNA sequencing, microarray comparative genomic hybridization, or MLPA (Gibriel & Adel, 2017). One advantage of this strategy is that any source of high quality genomic DNA, such as blood, can be used for the analysis. We are hopeful that with continued vigilance in sequencing patients and reporting mutations and associated treatment regimens, genotyping will not only diagnose CTE but also provide physicians the ability to better inform families for diagnosis, family planning, and ultimately cater specific treatment regimens.

8.2 | Lynch syndrome

Usually, *EPCAM*-associated Lynch syndrome is treated as classic Lynch syndrome with aggressive colorectal cancer surveillance, genetic counseling, and assessment of first-degree relatives. As described above, most *EPCAM*-associated Lynch syndrome mutations have a high risk of colorectal cancer and a low risk of endometrial cancer unless the *MSH2* gene is also affected (Dymerska et al., 2017; Kempers et al., 2011; Lynch et al., 2011; Perez-Cabornero et al., 2011). In theory, this suggests that women with *EPCAM*-associated Lynch syndrome

without deletions affecting *MSH2* should be focused primarily on colorectal cancers and may not need to undergo aggressive surveillance for endometrial cancer or prophylactic hysterectomy (Schmeler et al., 2006). In practice, general screening guidelines remain conservative and do not take the nuanced genetic difference into account.

Because *EPCAM*-associated Lynch syndrome is often caused by silencing of an otherwise functional *MSH2* gene, it is possible to envision that treatments that modify the *MSH2* promoter hypermethylation or prevent the read-through transcription of *EPCAM* and *MSH2* genes could preserve *MSH2* function and could act as a cancer prevention strategy. Such epigenetic approaches are being investigated for use in chemotherapy. For example, 5-azacytidine (Aza) and Zebularine are cytosine analogs that have been shown to cause reduced DNA methylation through degradation of a trapped drug-DNA methylase enzyme complex (Heerboth et al., 2014). Zebularine has been used to reactivate the silenced and hypermethylated *p16* tumor suppressor in a murine model (Cheng et al., 2003). Additionally, antisense targeting of the DNMT1 DNA methylase also can cause reexpression of a silenced *p16* gene and has been used in the treatment of an advanced renal cell carcinoma (Amato, 2007). Currently, DNA methylation inhibitors are being developed for chemotherapy and not for cancer suppression; careful analysis of any toxic side effects during long-term administration would need to be evaluated against any potential benefit in the suppression of tumor formation in Lynch syndrome caused by epigenetic silencing of *MSH2*.

Lynch syndrome is a disease in which a personalized medicine approach is feasible. Clinical history coupled with molecular tumor screening of a patient's colon carcinoma can lead to genetic testing and confirmation of Lynch syndrome in a patient and family members. Understanding of genetics, *EPCAM* associated or not, can directly inform the need for invasive and costly measures such as endometrial cancer screening and hysterectomy (Kempers et al., 2011). Further studies are needed to assess the impact, practicality, and cost-effectiveness of such approaches, given complex circumstances and treatment strategies (Guglielmo, Staropoli, Giancotti, & Mauro, 2018).

8.3 | *EPCAM* as a tumor marker and a cancer therapy target

EpCAM, the tumor antigen first identified with monoclonal antibodies (Herlyn et al., 1979), is expressed at high levels in many different adenocarcinomas and squamous cell carcinomas (Fang et al., 2017; Pan et al., 2018; Quak et al., 1990; Went et al., 2004), and has been suggested to be a cancer stem cell marker (Imrich, Hachmeister, & Gires, 2012). *EpCAM* has a role in promoting cell proliferation (Maetzel et al., 2009). *EpCAM* expression in a tumor is associated with poor prognosis in breast, ovarian, esophageal, and gallbladder cancer (Spizzo et al., 2004; Spizzo et al., 2006; Stoecklein et al., 2006; Trzpis et al., 2007; Varga et al., 2004) but is associated with somewhat improved prognosis in some colonic, gastric, and renal cancers (Seligson et al., 2004; Songun et al., 2005; Went et al., 2005; Went et al., 2006). *EpCAM* appears to promote cell proliferation in multiple studies involving both up- and downregulation of expression (Chaves-Perez et al., 2013; Maetzel et al., 2009; Munz et al., 2004; Wenqi et al., 2009),

which is due to the release of the intracellular C-terminal domain of the protein (EpiCD) by the tumor necrosis factor- α -converting enzyme and presenilin-2 proteases. The EpiCD fragment becomes localized to the nucleus and induces gene transcription effects that are oncogenic in immunodeficient mice (Maetzel et al., 2009). Overexpression of *EpCAM* has also been suggested to play a role in metastasis, due to a role of *EpCAM* in disrupting cell adhesion and promoting cell migration and cell invasion (Osta et al., 2004; Winter et al., 2003), which is consistent experiments in rats demonstrating that overexpression promotes metastases (Wurfel et al., 1999). As a molecule found at the surface of cancer cells, immunotherapies and antibody-drug conjugates targeting *EpCAM* have been investigated for treating cancers (Baeuerle & Gires, 2007; Lund et al., 2014; Miller et al., 2016; Moldenhauer et al., 2012), including catumaxomab, which has been approved for the treatment of malignant ascites from epithelial cancers (Linke, Klein, & Seimetz, 2010). Additionally, immunotherapies targeting *EpCAM* might also function to reduce the formation of metastases in some cancers of epithelial origin. Moreover, *EpCAM* processing is also an *EpCAM*-related cancer therapy target based on the fact that inhibition of the proteolysis inhibits growth signaling (Maetzel et al., 2009), and screening in an experimental setting has identified a number of potential inhibitors (Tretter et al., 2018). These observations suggest that adjuvant treatments targeting *EpCAM* biology may have the potential to broadly target epithelial tumors of different origins.

EpCAM's unique role in three disease states likely stems from its multifaceted role in depending on the tissue of interest and *EPCAM* variants and expression. Ongoing vigilance to report and correlate variants to clinical outcomes is necessary to inform understanding of disease progression and potential treatment strategies. Though the last two decades have led to numerous novel insights, many further studies are necessary to translate this knowledge to the bedside.

ACKNOWLEDGMENTS

We would like to thank the patients and families of those individuals with CTE for their support and participation. We thank Hal Hoffman, Lance Prince, Soumita Das, and Kim Barrett for useful discussions.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

ORCID

Mamata Sivagnanam  <https://orcid.org/0000-0002-3643-1326>

REFERENCES

- Al-Mayouf, S. M., Alswaied, N., Alkuraya, F. S., AlMehaidib, A., & Faqih, M. (2009). Tufting enteropathy and chronic arthritis: A newly recognized association with a novel *EpCAM* gene mutation. *Journal of Pediatric Gastroenterology and Nutrition*, 49(5), 642–644.
- AlMahamed, S., & Hammo, A. (2017). New mutations of *EpCAM* gene for tufting enteropathy in Saudi Arabia. *Saudi Journal of Gastroenterology: Official Journal of the Saudi Gastroenterology Association*, 23(2), 123–126.

- Amato, R. J. (2007). Inhibition of DNA methylation by antisense oligonucleotide MG98 as cancer therapy. *Clinical Genitourinary Cancer*, 5(7), 422–426. <https://doi.org/10.3816/CGC.2007.n.029>
- Baeuerle, P. A., & Gires, O. (2007). EpCAM (CD326) finding its role in cancer. *British Journal of Cancer*, 96(3), 417–423. <https://doi.org/10.1038/sj.bjc.6603494>
- Balzar, M., Prins, F. A., Bakker, H. A., Fleuren, G. J., Warnaar, S. O., & Litvinov, S. V. (1999). The structural analysis of adhesions mediated by Ep-CAM. *Experimental Cell Research*, 246(1), 108–121.
- Balzar, M., Winter, M., De Boer, C., & Litvinov, S. (1999). The biology of the 17–1A antigen (Ep-CAM). *Journal of Molecular Medicine*, 77(10), 699–712.
- Basak, S., Speicher, D., Eck, S., Wunner, W., Maul, G., Simmons, M. S., & Herlyn, D. (1998). Colorectal carcinoma invasion inhibition by CO17-1A/GA733 antigen and its murine homologue. *Journal of the National Cancer Institute*, 90(9), 691–697.
- Beck, N. S., Kang, I. S., & Suh, Y. L. (2001). Protracted diarrhea: Results of the five-year survey in a tertiary hospital in Korea. *Journal of Korean Medical Science*, 16(6), 736–741.
- Benchaouir, R., Robin, V., & Goyenvalle, A. (2015). Gene and splicing therapies for neuromuscular diseases. *Frontiers in Bioscience (Landmark Edition)*, 20, 1190–1233.
- Bergsagel, P. L., Victor-Kobrin, C., Timblin, C. R., Trepel, J., & Kuehl, W. M. (1992). A murine cDNA encodes a pan-epithelial glycoprotein that is also expressed on plasma cells. *Journal of Immunology*, 148(2), 590–596.
- Bodian, D. L., Vilboux, T., Hourigan, S. K., Jenevein, C. L., Mani, H., Kent, K. C., ... Hauser, N. S. (2017). Genomic analysis of an infant with intractable diarrhea and dilated cardiomyopathy. *Molecular Case Studies*, 3(6). <https://doi.org/10.1101/mcs.a002055>
- Chan, T. L., Yuen, S. T., Kong, C. K., Chan, Y. W., Chan, A. S., Ng, W. F., ... Leung, S. Y. (2006). Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nature Genetics*, 38(10), 1178–1183. <https://doi.org/10.1038/ng1866>
- Chaves-Perez, A., Mack, B., Maetzel, D., Kremling, H., Eggert, C., Harreus, U., & Gires, O. (2013). EpCAM regulates cell cycle progression via control of cyclin D1 expression. *Oncogene*, 32(5), 641–650. <https://doi.org/10.1038/ncr.2012.75>
- Cheng, J. C., Matsen, C. B., Gonzales, F. A., Ye, W., Greer, S., Marquez, V. E., ... Selker, E. U. (2003). Inhibition of DNA methylation and reactivation of silenced genes by zebularine. *Journal of the National Cancer Institute*, 95(5), 399–409.
- Cirulli, V., Ricordi, C., & Hayek, A. (1995). E-cadherin, NCAM, and EpCAM expression in human fetal pancreata. *Transplantation proceedings*, 27(6), 3335.
- Cutz, E., Rhoads, J. M., Drumm, B., Sherman, P. M., Durie, P. R., & Forstner, G. G. (1989). Microvillus inclusion disease: An inherited defect of brush-border assembly and differentiation. *New England Journal of Medicine*, 320(10), 646–651.
- d'Apolito, M., Pisanelli, D., Faletra, F., Giardino, I., Gigante, M., Pettoello-Mantovani, M., ... Campanozzi, A. (2016). Genetic analysis of Italian patients with congenital tufting enteropathy. *World Journal of Pediatrics*, 12(2), 219–224.
- Deininger, P. (2011). Alu elements: Know the SINEs. *Genome Biology*, 12(12), 236. <https://doi.org/10.1186/gb-2011-12-12-236>
- Dymerska, D., Golebiewska, K., Kuswik, M., Rudnicka, H., Scott, R. J., Billings, R., ... Kurzawski, G. (2017). New EPCAM founder deletion in Polish population. *Clinical Genetics*, 92(6), 649–653. <https://doi.org/10.1111/cge.13026>
- Eguchi, H., Kumamoto, K., Suzuki, O., Kohda, M., Tada, Y., Okazaki, Y., & Ishida, H. (2016). Identification of a Japanese Lynch syndrome patient with large deletion in the 3' region of the EPCAM gene. *Japanese Journal of Clinical Oncology*, 46(2), 178–184. <https://doi.org/10.1093/jjco/hyv172>
- Fang, S., Tian, H., Li, X., Jin, D., Li, X., Kong, J., ... Luo, Y. (2017). Clinical application of a microfluidic chip for immunocapture and quantification of circulating exosomes to assist breast cancer diagnosis and molecular classification. *PLoS One*, 12(4), e0175050.
- Fishel, R., Lescoe, M. K., Rao, M. R., Copeland, N. G., Jenkins, N. A., Garber, J., ... Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, 75(5), 1027–1038.
- Fisher, R. A. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.
- Fornaro, M., Dell'Arciprete, R., Stella, M., Bucci, C., Nutini, M., Capri, M. G., & Alberti, S. (1995). Cloning of the gene encoding Trop-2, a cell-surface glycoprotein expressed by human carcinomas. *International Journal of Cancer*, 62(5), 610–618.
- Furth, E. E., Li, J., Purev, E., Solomon, A. C., Rogler, G., Mick, R., ... Herlyn, D. (2006). Serum antibodies to EpCAM in healthy donors but not ulcerative colitis patients. *Cancer Immunology, Immunotherapy*, 55(5), 528–537. <https://doi.org/10.1007/s00262-005-0026-5>
- Gaber, A., Kim, S. J., Kaake, R. M., Bencina, M., Krogan, N., Sali, A., ... Lenarcic, B. (2018). EpCAM homo-oligomerization is not the basis for its role in cell-cell adhesion. *Scientific Reports*, 8(1), 13269. <https://doi.org/10.1038/s41598-018-31482-7>
- Gaiser, M. R., Lammermann, T., Feng, X., Igyarto, B. Z., Kaplan, D. H., Tessarollo, L., ... Udey, M. C. (2012). Cancer-associated epithelial cell adhesion molecule (EpCAM; CD326) enables epidermal Langerhans cell motility and migration in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 109(15), E889–E897. <https://doi.org/10.1073/pnas.1117674109>
- Gazzoli, I., Loda, M., Garber, J., Syngal, S., & Kolodner, R. D. (2002). A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Research*, 62(14), 3925–3928.
- Giardiello, F. M., Allen, J. I., Axilbund, J. E., Boland, C. R., Burke, C. A., Burt, R. W., ... Kaltenbach, T. (2014). Guidelines on genetic evaluation and management of Lynch syndrome: A consensus statement by the US Multi-Society Task Force on colorectal cancer. *Gastrointestinal Endoscopy*, 80(2), 197–220.
- Gibriel, A. A., & Adel, O. (2017). Advances in ligase chain reaction and ligation-based amplifications for genotyping assays: Detection and applications. *Mutation Research*, 773, 66–90. <https://doi.org/10.1016/j.mrrev.2017.05.001>
- Gonzalez, B., Denzel, S., Mack, B., Conrad, M., & Gires, O. (2009). EpCAM is involved in maintenance of the murine embryonic stem cell phenotype. *Stem Cells*, 27(8), 1782–1791. <https://doi.org/10.1002/stem.97>
- Goulet, O., Salomon, J., Ruemmele, F., de Serres, N. P.-M., & Brousse, N. (2007). Intestinal epithelial dysplasia (tufting enteropathy). *Orphanet Journal of Rare Diseases*, 2(1), 20.
- Guarinos, C., Castillejo, A., Barbera, V. M., Perez-Carbonell, L., Sanchez-Heras, A. B., Segura, A., ... Soto, J. L. (2010). EPCAM germ line deletions as causes of Lynch syndrome in Spanish patients. *The Journal of Molecular Diagnostics*, 12(6), 765–770. <https://doi.org/10.2353/jmoldx.2010.100039>
- Guerra, E., Lattanzio, R., La Sorda, R., Dini, F., Tiboni, G. M., Piantelli, M., & Alberti, S. (2012). mTrop1/Epcam knockout mice develop congenital tufting enteropathy through dysregulation of intestinal E-cadherin/ β -catenin. *PLoS One*, 7(11), e49302.
- Guglielmo, A., Staropoli, N., Giancotti, M., & Mauro, M. (2018). Personalized medicine in colorectal cancer diagnosis and treatment: A systematic

- review of health economic evaluations. *Cost Effectiveness and Resource Allocation*, 16, 2. <https://doi.org/10.1186/s12962-018-0085-z>
- Heerboth, S., Lapinska, K., Snyder, N., Leary, M., Rollinson, S., & Sarkar, S. (2014). Use of epigenetic drugs in disease: An overview. *Genetics & epigenetics*, 6, 9–19. <https://doi.org/10.4137/geg.s12270>
- Hemminki, A., Peltomäki, P., Mecklin, J. P., Jarvinen, H., Salovaara, R., Nystrom-Lahti, M., ... Aaltonen, L. A. (1994). Loss of the wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. *Nature Genetics*, 8(4), 405–410. <https://doi.org/10.1038/ng1294-405>
- Herlyn, M., Steplewski, Z., Herlyn, D., & Koprowski, H. (1979). Colorectal carcinoma-specific antigen: Detection by means of monoclonal antibodies. *Proceedings of the National Academy of Sciences of the United States of America*, 76(3), 1438–1442.
- Hitchins, M. P., & Burn, J. (2011). Alu in Lynch syndrome: A danger SINE? *Cancer Prevention Research*, 4(10), 1527–1530. <https://doi.org/10.1158/1940-6207.capr-11-0417>
- Hong, S. N., Dunn, J. C., Stelzner, M., & Martin, M. G. (2017). Concise review: The potential use of intestinal stem cells to treat patients with intestinal failure. *Stem Cells Translational Medicine*, 6(2), 666–676. <https://doi.org/10.5966/sctm.2016-0153>
- Hua, Y., Sahashi, K., Hung, G., Rigo, F., Passini, M. A., Bennett, C. F., & Krainer, A. R. (2010). Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes & Development*, 24(15), 1634–1644.
- Hua, Y., Vickers, T. A., Okunola, H. L., Bennett, C. F., & Krainer, A. R. (2008). Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *The American Journal of Human Genetics*, 82(4), 834–848.
- Huth, C., Kloor, M., Voigt, A. Y., Bozukova, G., Evers, C., Gaspar, H., ... Blaker, H. (2012). The molecular basis of EPCAM expression loss in Lynch syndrome-associated tumors. *Modern Pathology*, 25(6), 911–916. <https://doi.org/10.1038/modpathol.2012.30>
- Imrich, S., Hachmeister, M., & Gires, O. (2012). EpCAM and its potential role in tumor-initiating cells. *Cell Adhesion & Migration*, 6(1), 30–38.
- Kang, S. Y., Park, C. K., Chang, D. K., Kim, J. W., Son, H. J., Cho, Y. B., ... Kim, K. M. (2015). Lynch-like syndrome: Characterization and comparison with EPCAM deletion carriers. *International Journal of Cancer*, 136(7), 1568–1578. <https://doi.org/10.1002/ijc.29133>
- Kempers, M. J., Kuiper, R. P., Ockeloen, C. W., Chappuis, P. O., Hutter, P., Rahner, N., ... Morak, M. (2011). Risk of colorectal and endometrial cancers in EPCAM deletion-positive Lynch syndrome: A cohort study. *The Lancet Oncology*, 12(1), 49–55.
- Kloor, M., Voigt, A. Y., Schackert, H. K., Schirmacher, P., von Knebel Doeberitz, M., & Blaker, H. (2011). Analysis of EPCAM protein expression in diagnostics of Lynch syndrome. *Journal of Clinical Oncology*, 29(2), 223–227. <https://doi.org/10.1200/jco.2010.32.0820>
- Ko, J. S., Seo, J. K., Shim, J. O., Hwang, S. H., Park, H. S., & Kang, G. H. (2010). Tufting enteropathy with EpCAM mutations in two siblings. *Gut and Liver*, 4(3), 407–410.
- Kole, R., Krainer, A. R., & Altman, S. (2012). RNA therapeutics: Beyond RNA interference and antisense oligonucleotides. *Nature Reviews. Drug Discovery*, 11(2), 125–140.
- Kolodner, R. D. (2016). A personal historical view of DNA mismatch repair with an emphasis on eukaryotic DNA mismatch repair. *DNA Repair*, 38, 3–13. <https://doi.org/10.1016/j.dnarep.2015.11.009>
- Kovacs, M. E., Papp, J., Szentirmay, Z., Otto, S., & Olah, E. (2009). Deletions removing the last exon of TACSTD1 constitute a distinct class of mutations predisposing to Lynch syndrome. *Human Mutation*, 30(2), 197–203. <https://doi.org/10.1002/humu.20942>
- Kozan, P. A., McGeough, M. D., Peña, C. A., Mueller, J. L., Barrett, K. E., Marchelletta, R. R., & Sivagnanam, M. (2015). Mutation of EpCAM leads to intestinal barrier and ion transport dysfunction. *Journal of Molecular Medicine*, 93(5), 535–545.
- Kuiper, R. P., Vissers, L. E., Venkatachalam, R., Bodmer, D., Hoenselaar, E., Goossens, M., ... Ligtenberg, M. J. (2011). Recurrence and variability of germline EPCAM deletions in Lynch syndrome. *Human Mutation*, 32(4), 407–414. <https://doi.org/10.1002/humu.21446>
- Ladwein, M., Pape, U.-F., Schmidt, D.-S., Schnölzer, M., Fiedler, S., Langbein, L., ... Zöller, M. (2005). The cell-cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7. *Experimental Cell Research*, 309(2), 345–357.
- Lei, Z., Maeda, T., Tamura, A., Nakamura, T., Yamazaki, Y., Shiratori, H., ... Hamada, H. (2012). EpCAM contributes to formation of functional tight junction in the intestinal epithelium by recruiting claudin proteins. *Developmental Biology*, 371(2), 136–145. <https://doi.org/10.1016/j.ydbio.2012.07.005>
- Lemale, J., Coulomb, A., Dubern, B., Boudjemaa, S., Viola, S., Josset, P., ... Girardet, J.-P. (2011). Intractable diarrhea with tufting enteropathy: A favorable outcome is possible. *Journal of Pediatric Gastroenterology and Nutrition*, 52(6), 734–739.
- Li-Chang, H. H., Driman, D. K., Levin, H., Siu, V. M., Scanlan, N. L., Buckley, K., ... Ainsworth, P. J. (2013). Colorectal cancer in a 9-year-old due to combined EPCAM and MSH2 germline mutations: Case report of a unique genotype and immunophenotype. *Journal of Clinical Pathology*, 66(7), 631–633. <https://doi.org/10.1136/jclinpath-2012-201376>
- Ligtenberg, M. J., Kuiper, R. P., Chan, T. L., Goossens, M., Hebeda, K. M., Voorendt, M., ... Hendriks-Cornelissen, S. J. (2009). Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nature genetics*, 41(1), 112–117.
- Linke, R., Klein, A., & Seimetz, D. (2010). Catumaxomab: Clinical development and future directions. *MAbs*, 2(2), 129–136.
- Litvinov, S. V., Bakker, H. A., Gourevitch, M. M., Velders, M. P., & Warnaar, S. O. (1994). Evidence for a role of the epithelial glycoprotein 40 (EpCAM) in epithelial cell-cell adhesion. *Cell Adhesion and Communication*, 2(5), 417–428.
- Litvinov, S. V., Balzar, M., Winter, M. J., Bakker, H. A., Briare-de Bruijn, I. H., Prins, F., ... Warnaar, S. O. (1997). Epithelial cell adhesion molecule (EpCAM) modulates cell-cell interactions mediated by classic cadherins. *The Journal of cell biology*, 139(5), 1337–1348.
- Litvinov, S. V., Velders, M. P., Bakker, H. A., Fleuren, G. J., & Warnaar, S. O. (1994). Ep-CAM: A human epithelial antigen is a homophilic cell-cell adhesion molecule. *Journal of Cell Biology*, 125(2), 437–446.
- Lund, K., Bostad, M., Skarpen, E., Braunagel, M., Kiprijanov, S., Krauss, S., ... Selbo, P. K. (2014). The novel EpCAM-targeting monoclonal antibody 3–171 linked to saporin is highly cytotoxic after photochemical internalization in breast, pancreas and colon cancer cell lines. *MAbs*, 6(4), 1038–1050.
- Lynch, H. T., Lynch, P., Lanspa, S., Snyder, C., Lynch, J., & Boland, C. (2009). Review of the Lynch syndrome: History, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clinical genetics*, 76(1), 1–18.
- Lynch, H. T., Riegert-Johnson, D. L., Snyder, C., Lynch, J. F., Hagenkord, J., Boland, C. R., ... Davies, J. (2011). Lynch syndrome-associated extra-colonic tumors are rare in two extended families with the same EPCAM deletion. *The American Journal of Gastroenterology*, 106(10), 1829–1836.
- Maetzel, D., Denzel, S., Mack, B., Canis, M., Went, P., Benk, M., ... Gires, O. (2009). Nuclear signalling by tumour-associated antigen EpCAM. *Nature Cell Biology*, 11(2), 162–171. <https://doi.org/10.1038/ncb1824>

- Maquat, L. E. (2004). Nonsense-mediated mRNA decay: Splicing, translation and mRNP dynamics. *Nature Reviews Molecular Cell Biology*, 5(2), 89–99. <https://doi.org/10.1038/nrm1310>
- Miller, M. L., Fishkin, N. E., Li, W., Whiteman, K. R., Kovtun, Y., Reid, E. E., ... Chari, R. V. (2016). A new class of antibody-drug conjugates with potent DNA alkylating activity. *Molecular Cancer Therapeutics*, 15(8), 1870–1878. <https://doi.org/10.1158/1535-7163.mct-16-0184>
- Moldenhauer, G., Salnikov, A. V., Luttgau, S., Herr, I., Anderl, J., & Faulstich, H. (2012). Therapeutic potential of amanitin-conjugated anti-epithelial cell adhesion molecule monoclonal antibody against pancreatic carcinoma. *Journal of the National Cancer Institute*, 104(8), 622–634. <https://doi.org/10.1093/jnci/djs140>
- Mueller, J. L., McGeough, M. D., Pena, C. A., & Sivagnanam, M. (2014). Functional consequences of EpCam mutation in mice and men. *American Journal of Physiology. Gastrointestinal and Liver Physiology*, 306(4), G278–G288. <https://doi.org/10.1152/ajpgi.00286.2013>
- Munz, M., Kieu, C., Mack, B., Schmitt, B., Zeidler, R., & Gires, O. (2004). The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. *Oncogene*, 23(34), 5748–5758. <https://doi.org/10.1038/sj.onc.1207610>
- Mur, P., Pineda, M., Romero, A., Del Valle, J., Borrás, E., Canal, A., ... Capella, G. (2014). Identification of a founder EPCAM deletion in Spanish Lynch syndrome families. *Clinical Genetics*, 85(3), 260–266. <https://doi.org/10.1111/cge.12152>
- Musulen, E., Blanco, I., Carrato, C., Fernandez-Figueras, M. T., Pineda, M., Capella, G., & Ariza, A. (2013). Usefulness of epithelial cell adhesion molecule expression in the algorithmic approach to Lynch syndrome identification. *Human Pathology*, 44(3), 412–416. <https://doi.org/10.1016/j.humpath.2012.06.006>
- Nagao, K., Zhu, J., Heneghan, M. B., Hanson, J. C., Morasso, M. I., Tesarollo, L., ... Udey, M. C. (2009). Abnormal placental development and early embryonic lethality in EpCAM-null mice. *PLoS One*, 4(12), e8543. <https://doi.org/10.1371/journal.pone.0008543>
- Nagasaka, T., Rhee, J., Kloor, M., Gebert, J., Naomoto, Y., Boland, C. R., & Goel, A. (2010). Somatic hypermethylation of MSH2 is a frequent event in Lynch Syndrome colorectal cancers. *Cancer Research*, 70(8), 3098–3108. <https://doi.org/10.1158/0008-5472.can-09-3290>
- Niessen, R. C., Hofstra, R. M., Westers, H., Ligtenberg, M. J., Kooi, K., Jager, P. O., ... Sijmons, R. H. (2009). Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. *Genes, Chromosomes & Cancer*, 48(8), 737–744. <https://doi.org/10.1002/gcc.20678>
- Osta, W. A., Chen, Y., Mikhitarian, K., Mitas, M., Salem, M., Hannun, Y. A., ... Gillanders, W. E. (2004). EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. *Cancer Research*, 64(16), 5818–5824. <https://doi.org/10.1158/0008-5472.Can-04-0754>
- Pan, M., Schinke, H., Luxenburger, E., Kranz, G., Shakhtour, J., Libl, D., ... Lenarčič, B. (2018). EpCAM ectodomain EpEX is a ligand of EGFR that counteracts EGF-mediated epithelial-mesenchymal transition through modulation of phospho-ERK1/2 in head and neck cancers. *PLoS Biology*, 16(9), e2006624.
- Paramesh, A. S., Fishbein, T., Tschernia, A., Leleiko, N., Magid, M. S., Gondolesi, G. E., & Kaufman, S. S. (2003). Isolated small bowel transplantation for tufting enteropathy. *Journal of Pediatric Gastroenterology and Nutrition*, 36(1), 138–140.
- Pavsic, M., Guncar, G., Djinovic-Carugo, K., & Lenarcic, B. (2014). Crystal structure and its bearing towards an understanding of key biological functions of EpCAM. *Nature communications*, 5, 4764.
- Pêgas, K. L., Cambruzzi, E., Ferrelli, R. S., Silva, C. S. d., Guedes, R. R., Adami, M., ... Steinhaus, C. (2014). Tufting enteropathy with EpCAM mutation: Case report. *Jornal Brasileiro de Patologia e Medicina Laboratorial*, 50(3), 234–237.
- Perez-Cabornero, L., Infante Sanz, M., Velasco Sampedro, E., Lastra Aras, E., Acedo Becares, A., Miner Pino, C., & Duran Dominguez, M. (2011). Frequency of rearrangements in Lynch syndrome cases associated with MSH2: Characterization of a new deletion involving both EPCAM and the 5' part of MSH2. *Cancer Prevention Research*, 4(10), 1556–1562. <https://doi.org/10.1158/1940-6207.capr-11-0080>
- Pitman, E. J. (1937). Significance tests which may be applied to samples from any populations. *Supplement to the Journal of the Royal Statistical Society*, 4(1), 119–130.
- Pitman, E. J. G. (1938). Significance tests which may be applied to samples from any populations: III. The analysis of variance test. *Biometrika*, 29(3/4), 322–335.
- Quak, J. J., Van Dongen, G., Brakkee, J. G., Hayashida, D. J., Balm, A. J., Snow, G. B., & Meijer, C. J. (1990). Production of a monoclonal antibody (K 931) to a squamous cell carcinoma associated antigen identified as the 17-1A antigen. *Hybridoma*, 9(4), 377–387. <https://doi.org/10.1089/hyb.1990.9.377>
- Ranganathan, S., Schmitt, L. A., & Sindhi, R. (2014). Tufting enteropathy revisited: The utility of MOC31 (EpCAM) immunohistochemistry in diagnosis. *The American journal of surgical pathology*, 38(2), 265–272.
- Reifen, R. M., Cutz, E., Griffiths, A.-M., Ngan, B. Y., & Sherman, P. M. (1994). Tufting enteropathy: A newly recognized clinicopathological entity associated with refractory diarrhea in infants. *Journal of Pediatric Gastroenterology and Nutrition*, 18(3), 379–385.
- Ren, Z., Lin, P. Y., Klintworth, G. K., Iwata, F., Munier, F. L., Schorderet, D. F., ... Hejtmancik, J. F. (2002). Allelic and locus heterogeneity in autosomal recessive gelatinous drop-like corneal dystrophy. *Human Genetics*, 110(6), 568–577. <https://doi.org/10.1007/s00439-002-0729-z>
- Rossi, B. M., Palmero, E. I., Lopez-Kostner, F., Sarroca, C., Vaccaro, C. A., Spiranelli, F., ... Dominguez-Valentin, M. (2017). A survey of the clinicopathological and molecular characteristics of patients with suspected Lynch syndrome in Latin America. *BMC Cancer[Electronic Resource]*, 17(1), 623. <https://doi.org/10.1186/s12885-017-3599-4>
- Rumilla, K., Schowalter, K. V., Lindor, N. M., Thomas, B. C., Mensink, K. A., Gallinger, S., ... Thibodeau, S. N. (2011). Frequency of deletions of EPCAM (TACSTD1) in MSH2-associated Lynch syndrome cases. *The Journal of Molecular Diagnostics*, 13(1), 93–99. <https://doi.org/10.1016/j.jmoldx.2010.11.011>
- Salomon, J., Espinosa-Parrilla, Y., Goulet, O., Guigue, P., Canioni, D., Bruneau, J., ... Jeanpierre, M. (2011). A founder effect at the EPCAM locus in congenital tufting enteropathy in the Arabic Gulf. *European Journal of Medical Genetics*, 54(3), 319–322.
- Salomon, J., Goulet, O., Canioni, D., Brousse, N., Lemale, J., Tounian, P., ... Ruemmele, F. (2014). Genetic characterization of congenital tufting enteropathy: Epcam associated phenotype and involvement of SPINT2 in the syndromic form. *Human genetics*, 133(3), 299–310.
- Sarrach, S., Huang, Y., Niedermeyer, S., Hachmeister, M., Fischer, L., Gille, S., ... Gires, O. (2018). Spatiotemporal patterning of EpCAM is important for murine embryonic endo- and mesodermal differentiation. *Scientific Reports*, 8(1), 1801. <https://doi.org/10.1038/s41598-018-20131-8>
- Schmeler, K. M., Lynch, H. T., Chen, L.-m., Munsell, M. F., Soliman, P. T., Clark, M. B., ... Conrad, P. G. (2006). Prophylactic surgery to reduce the risk of gynecologic cancers in the Lynch syndrome. *New England Journal of Medicine*, 354(3), 261–269.
- Schnell, U., Kuipers, J., Mueller, J. L., Veenstra-Algra, A., Sivagnanam, M., & Giepmans, B. N. (2013). Absence of cell-surface EpCAM in congenital tufting enteropathy. *Human molecular genetics*, 22(13), 2566–2571.
- Sekine, S., Ogawa, R., Saito, S., Ushiyama, M., Shida, D., Nakajima, T., ... Sugano, K. (2017). Cytoplasmic MSH2 immunoreactivity in a patient with Lynch syndrome with an EPCAM-MSH2 fusion. *Histopathology*, 70(4), 664–669. <https://doi.org/10.1111/his.13104>

- Seligson, D. B., Pantuck, A. J., Liu, X., Huang, Y., Horvath, S., Bui, M. H., ... Figlin, R. A. (2004). Epithelial cell adhesion molecule (KSA) expression: Pathobiology and its role as an independent predictor of survival in renal cell carcinoma. *Clinical Cancer Research*, 10(8), 2659–2669.
- Shakhnovich, V., Dinwiddie, D., Hildreth, A., Attard, T., & Kingsmore, S. (2017). A novel compound-heterozygous EpCAM mutation in tufting enteropathy. *Journal of Pediatric Gastroenterology and Nutrition*, 64(1), e14–e16.
- Sherman, P. M., Mitchell, D. J., & Cutz, E. (2004). Neonatal enteropathies: Defining the causes of protracted diarrhea of infancy. *Journal of Pediatric Gastroenterology and Nutrition*, 38(1), 16–26.
- Shlien, A., Campbell, B. B., de Borja, R., Alexandrov, L. B., Merico, D., Wedge, D., ... Tabori, U. (2015). Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nature Genetics*, 47(3), 257–262. <https://doi.org/10.1038/ng.3202>
- Sivagnanam, M., Mueller, J. L., Lee, H., Chen, Z., Nelson, S. F., Turner, D., ... Libiger, O. (2008). Identification of EpCAM as the gene for congenital tufting enteropathy. *Gastroenterology*, 135(2), 429–437.
- Sivagnanam, M., Schaible, T., Szigeti, R., Byrd, R. H., Finegold, M. J., Ranganathan, S., ... Kellermayer, R. (2010). Further evidence for EpCAM as the gene for congenital tufting enteropathy. *American Journal of Medical Genetics Part A*, 152(1), 222–224.
- Slanchev, K., Carney, T. J., Stemmler, M. P., Koschorz, B., Amsterdam, A., Schwarz, H., & Hammerschmidt, M. (2009). The epithelial cell adhesion molecule EpCAM is required for epithelial morphogenesis and integrity during zebrafish epiboly and skin development. *PLoS Genetics*, 5(7), e1000563. <https://doi.org/10.1371/journal.pgen.1000563>
- Smit, A. F. A., Hubley, R., & Green, P. (2013–2015). RepeatMasker Open-4.0.
- Sokal, R. R., & Rohlf, F. J. (1994). *Biometry: The principles and practice of statistics in biological research*. New York, NY: W.H. Freeman
- Songun, I., Litvinov, S. V., van de Velde, C. J., Pals, S. T., Hermans, J., & van Krieken, J. H. (2005). Loss of Ep-CAM (CO17-1A) expression predicts survival in patients with gastric cancer. *British Journal of Cancer*, 92(9), 1767–1772. <https://doi.org/10.1038/sj.bjc.6602519>
- Spaepen, M., Neven, E., Sagaert, X., De Hertogh, G., Beert, E., Wimmer, K., ... Brems, H. (2013). EPCAM germline and somatic rearrangements in Lynch syndrome: Identification of a novel 3'EPCAM deletion. *Genes, Chromosomes & Cancer*, 52(9), 845–854. <https://doi.org/10.1002/gcc.22080>
- Spizzo, G., Gastl, G., Obrist, P., Went, P., Dirnhofer, S., Bischoff, S., ... Stopatschinskaya, S. (2004). High Ep-CAM expression is associated with poor prognosis in node-positive breast cancer. *Breast cancer research and treatment*, 86(3), 207–213.
- Spizzo, G., Went, P., Dirnhofer, S., Obrist, P., Moch, H., Baeuerle, P. A., ... Zeimet, A. G. (2006). Overexpression of epithelial cell adhesion molecule (Ep-CAM) is an independent prognostic marker for reduced survival of patients with epithelial ovarian cancer. *Gynecologic Oncology*, 103(2), 483–488. <https://doi.org/10.1016/j.ygyno.2006.03.035>
- Stoecklein, N. H., Siegmund, A., Scheunemann, P., Luebke, A. M., Erbersdobler, A., Verde, P. E., ... Hosch, S. B. (2006). Ep-CAM expression in squamous cell carcinoma of the esophagus: A potential therapeutic target and prognostic marker. *BMC Cancer [Electronic Resource]*, 6, 165. <https://doi.org/10.1186/1471-2407-6-165>
- Tang, W., Huang, T., Xu, Z., & Huang, Y. (2018). Novel mutations in EPCAM cause congenital tufting enteropathy. *Journal of clinical gastroenterology*, 52(1), e1–e6.
- Thiagarajah, J. R., Kamin, D. S., Acra, S., Goldsmith, J. D., Roland, J. T., Lencer, W. I., ... Martin, M. G. (2018). Advances in evaluation of chronic diarrhea in infants. *Gastroenterology*, 154(8), 2045–2059.e6. <https://doi.org/10.1053/j.gastro.2018.03.067>
- Thoeni, C., Amir, A., Guo, C., Zhang, S., Avitzur, Y., Heng, Y., ... Muise, A. M. (2014). A novel nonsense mutation in the EpCAM gene in a patient with congenital tufting enteropathy. *Journal of Pediatric Gastroenterology and Nutrition*, 58(1), 18–21.
- Treetipsatit, J., & Hazard, F. K. (2014). Features of gastric and colonic mucosa in congenital enteropathies: A study in histology and immunohistochemistry. *The American Journal of Surgical Pathology*, 38(12), 1697–1706.
- Tretter, J. Y., Schorpp, K., Luxenburger, E., Trambauer, J., Steiner, H., Hadian, K., ... Niessing, D. (2018). A high-content screen for small-molecule regulators of epithelial cell-adhesion molecule (EpCAM) cleavage yields a robust inhibitor. *Journal of Biological Chemistry*, 293(23), 8994–9005. <https://doi.org/10.1074/jbc.RA118.002776>
- Trzpis, M., McLaughlin, P. M., de Leij, L. M., & Harmsen, M. C. (2007). Epithelial cell adhesion molecule: More than a carcinoma marker and adhesion molecule. *The American journal of pathology*, 171(2), 386–395.
- Trzpis, M., McLaughlin, P. M., van Goor, H., Brinker, M. G., van Dam, G. M., de Leij, L. M., ... Harmsen, M. C. (2008). Expression of EpCAM is up-regulated during regeneration of renal epithelia. *Journal of Pathology*, 216(2), 201–208. <https://doi.org/10.1002/path.2396>
- Tsaktanis, T., Kremling, H., Pavsic, M., von Stackelberg, R., Mack, B., Fukumori, A., ... Huang, Z. (2015). Cleavage and cell adhesion properties of human epithelial cell adhesion molecule (HEPCAM). *Journal of Biological Chemistry*, 290(40), 24574–24591.
- Tsujikawa, M., Kurahashi, H., Tanaka, T., Nishida, K., Shimomura, Y., Tano, Y., & Nakamura, Y. (1999). Identification of the gene responsible for gelatinous drop-like corneal dystrophy. *Nature Genetics*, 21(4), 420–423. <https://doi.org/10.1038/7759>
- van der Klift, H., Wijnen, J., Wagner, A., Verkuilen, P., Tops, C., Otway, R., ... Fodde, R. (2005). Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible for hereditary nonpolyposis colorectal cancer (HNPCC). *Genes, Chromosomes & Cancer*, 44(2), 123–138. <https://doi.org/10.1002/gcc.20219>
- Varga, M., Obrist, P., Schneeberger, S., Muhlmann, G., Felgel-Farnholz, C., Fong, D., ... Spizzo, G. (2004). Overexpression of epithelial cell adhesion molecule antigen in gallbladder carcinoma is an independent marker for poor survival. *Clinical Cancer Research*, 10(9), 3131–3136.
- Vercollone, J. R., Balzar, M., Litvinov, S. V., Yang, W., & Cirulli, V. (2015). MMTV/LTR promoter-driven transgenic expression of EpCAM leads to the development of large pancreatic islets. *Journal of Histochemistry & Cytochemistry*, 63(8), 613–625.
- Verloes, A., Lombet, J., Lambert, Y., Hubert, A. F., Deprez, M., Fridman, V., ... Sokal, E. (1997). Tricho-hepato-enteric syndrome: Further delineation of a distinct syndrome with neonatal hemochromatosis phenotype, intractable diarrhea, and hair anomalies. *American Journal of Medical Genetics Part A*, 68(4), 391–395.
- Villablanca, E. J., Renucci, A., Sapede, D., Lec, V., Soubiran, F., Sandoval, P. C., ... Allende, M. L. (2006). Control of cell migration in the zebrafish lateral line: Implication of the gene "tumour-associated calcium signal transducer," *tacstd*. *Developmental Dynamics*, 235(6), 1578–1588. <https://doi.org/10.1002/dvdy.20743>
- Wenqi, D., Li, W., Shanshan, C., Bei, C., Yafei, Z., Feihu, B., ... Daiming, F. (2009). EpCAM is overexpressed in gastric cancer and its downregulation suppresses proliferation of gastric cancer. *Journal of Cancer Research and Clinical Oncology*, 135(9), 1277–1285. <https://doi.org/10.1007/s00432-009-0569-5>
- Went, P., Dirnhofer, S., Salvisberg, T., Amin, M. B., Lim, S. D., Diener, P. A., & Moch, H. (2005). Expression of epithelial cell adhesion molecule (EpCam) in renal epithelial tumors. *American Journal of Surgical Pathology*, 29(1), 83–88.

- Went, P., Vasei, M., Bubendorf, L., Terracciano, L., Tornillo, L., Riede, U., ... Baeuerle, P. A. (2006). Frequent high-level expression of the immunotherapeutic target Ep-CAM in colon, stomach, prostate and lung cancers. *British Journal of Cancer*, 94(1), 128–135. <https://doi.org/10.1038/sj.bjc.6602924>
- Went, P. T., Lugli, A., Meier, S., Bundi, M., Mirlacher, M., Sauter, G., & Dirnhofer, S. (2004). Frequent EpCam protein expression in human carcinomas. *Human Pathology*, 35(1), 122–128.
- Wimmer, K., & Etzler, J. (2008). Constitutional mismatch repair-deficiency syndrome: Have we so far seen only the tip of an iceberg? *Human Genetics*, 124(2), 105–122. <https://doi.org/10.1007/s00439-008-0542-4>
- Winter, M. J., Nagelkerken, B., Mertens, A. E., Rees-Bakker, H. A., Briare-de Bruijn, I. H., & Litvinov, S. V. (2003). Expression of Ep-CAM shifts the state of cadherin-mediated adhesions from strong to weak. *Experimental Cell Research*, 285(1), 50–58.
- Wu, C. J., Feng, X., Lu, M., Morimura, S., & Udey, M. C. (2017). Matriptase-mediated cleavage of EpCAM destabilizes claudins and dysregulates intestinal epithelial homeostasis. *Journal of Clinical Investigation*, 127(2), 623–634. <https://doi.org/10.1172/JCI88428>
- Wu, C. J., Mannan, P., Lu, M., & Udey, M. C. (2013). Epithelial cell adhesion molecule (EpCAM) regulates claudin dynamics and tight junctions. *Journal of Biological Chemistry*, 288(17), 12253–12268. <https://doi.org/10.1074/jbc.M113.457499>
- Wurfel, J., Rosel, M., Seiter, S., Claas, C., Herlevsen, M., Weth, R., & Zoller, M. (1999). Metastasis-association of the rat ortholog of the human epithelial glycoprotein antigen EGP314. *Oncogene*, 18(14), 2323–2334. <https://doi.org/10.1038/sj.onc.1202542>
- Zeinalian, M., Hashemzadeh-Chaleshtori, M., Salehi, R., & Emami, M. H. (2018). Clinical aspects of microsatellite instability testing in colorectal cancer. *Advanced Biomedical Research*, 7, 28. https://doi.org/10.4103/abr.abr_185_16

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Pathak SJ, Mueller JL, Okamoto K, et al. EPCAM mutation update: Variants associated with congenital tufting enteropathy and Lynch syndrome. *Human Mutation*. 2019;40:142–161. <https://doi.org/10.1002/humu.23688>